The Object Inside: Assessing 3D Examination with a Spherical Handheld Perspective-Corrected Display

Francois Berard, Thibault Louis
Univ. Grenoble Alpes, CNRS, LIG
F-38000 Grenoble France
francois.berard,thibault.louis@imag.fr

ABSTRACT

Handheld Perspective Corrected Displays (HPCDs) can create the feeling of holding a virtual 3D object. They offer a direct interaction that is isomorphic to the manipulation of physical objects. This illusion depends on the ability to provide a natural visuomotor coupling. High performances systems are thus required to evaluate the fundamental merits of HPCDs. We built a spherical HPCD using external projection. The system offers a lightweight wireless seamless display with head-coupled stereo, robust tracking, and low latency. We compared users' performances with this HPCD and two other interactions that used a fixed planar display and either a touchpad or the spherical display as an indirect input. The task involved the inspection of complex virtual 3D puzzles. Physical puzzles were also tested as references. Contrary to expectations, all virtual interactions were found to be more efficient than a more "natural" physical puzzle. The HPCD yielded lower performances than the touchpad. This study indicates that the object examination task did not benefit from the accurate and precise rotations offered by the HPCD, but benefited from the high C/D gain of the touchpad.

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces - Input devices and strategies; I.3.7 Computer Graphics: Three-Dimensional Graphics and Realism - Virtual reality

Author Keywords

handheld perspective corrected display (HPCD); 3D display; object examination; depth perception; isomorphic rotation; evaluation

INTRODUCTION

When we need to get a detailed understanding of the structure of a complex object such as the internals of a mechanical watch, we tend to grab the object with one or two hands and orient it in many different ways. This ability to quickly get radically different vantage points allows an efficient mental reconstruction of the object. Moreover, the active aspect of this manipulation seems important: passively watching an animation may not be as efficient as actively controlling the point of view [8, 13].

Figure 1. The three methods used to provide visual stimuli in the experiment: on a planar screen (left), on the spherical HPCD (center), and using a physical object (right). All three provided head-coupled stereo.

In the digital world, getting a detailed understanding of complex 3D structures, such as 3D models or complex visualizations, appears to be more difficult because virtual objects don't have a physical existence and thus they cannot be directly touched and rotated in our hands. This problem can be mitigated by the Handheld Perspective Corrected Display (HPCD) approach introduced by Stavness et al. [16]. HPCDs are small handheld displays that have a volume (i.e. they are non-planar). By tracking the position of the display and the user's head, and by showing the corresponding perspective images on the display's surface, HPCDs create the illusion that the virtual object is inside the display, as illustrated on Figure 1, center. When the virtual object is rigidly "attached" inside a HPCD, its rotations are controlled by the intuitive rotations of the display itself. Although HPCDs don't have the exact shape of the virtual object that they contain, previous research indicate that this does not affect users control of rotations [20]. The same study, however, showed that efficient rotations required to have the hand physically in the same location as the virtual object being manipulated, a property satisfied by HPCDs. Hence, the HPCD approach appears as promising for an efficient examination of complex 3D objects.

Our main contribution is an empirical study aimed at assessing the benefits of HPCDs for object examination. We compared HPCD interaction with two interactions using a standard planar display for output, and either a touchpad or the spherical display as indirect input. We also tested the interaction with physical objects as a baseline. The secondary contribution is technical: we improved the state of the art in terms of HPCD

CHI 2017, May 06 - 11, 2017, Denver, CO, USA

ACM 978-1-4503-4655-9/17/05 ...$15.00
DOI: http://dx.doi.org/10.1145/3025453.3025806

interaction, with the goal that the experiment results depict the true merits of the HPCD approach, and not the limits of the experimental system. We present a novel approach using video projection. It allows the implementation of a lightweight (~0.1 kg) 14 cm wireless *spherical* HPCD. It provides head-coupled stereo rendering with no seams and low latency (35 ms).

In the following, we start with a review of related work on depth perception, 3D rotations and object examination. After this, we present the design and implementation of the high performance spherical HPCD. We then present our experimental evaluation of object inspection. We finish with a general discussion of the experiment results and the potential of HPCDs, and conclude.

RELATED WORK

Depth Perception with Virtual Objects

Depth perception can serve many purposes, such as estimating the absolute distance between objects, or estimating if two moving objects will collide [6]. Here, we focus on depth perception used in the examination of a complex unknown 3D object. A continuous stream of efforts has improved the way 3D scenes are created and provided to users [1, 6, 18, 19]. Current systems generate realistic 3D scenes using various combinations of depth cues similar to those of the physical world: perspective projection, shading, stereo, and motion parallax; which are all supported by our system. Complex object examination was extensively studied using path-tracing tasks: users are shown two nodes on a tree or a graph and must tell if the two nodes are connected [1, 6, 19, 18]. The total number of nodes in the graph was used to control a *static* complexity of the graph. A more *dynamic* complexity of the *task* can be controlled by increasing the number of nodes that must be traversed to provide the answer, although only limited numbers of steps were tested (up to 4). Previous works tested different depth cues in order to assess their effect on depth perception. Varying the viewpoint on the object was found more important than stereo [18], but a combination of both stereo and head coupling was consistently found as the most effective [6, 18, 19], we used both in our experiment. We also introduce a novel task, which affords a path-tracing behavior that is more continuous than when connecting nodes in a graph, with the goal of fostering object manipulation.

3D Rotations of Virtual Objects

The human dexterity in manipulating physical objects is not directly reproducible in the virtual world. Various approach have been studied using indirect interaction [2, 11], direct and indirect tangible “props” [5, 10, 11, 20], multi-touch on planar surfaces [7], or contactless hand motions [9]. In the specific case of the control of 3D rotations, empirical studies where conducted using orientation matching tasks: participants had to match the orientation of a controlled object to that of a model shown by the system [5, 11, 14, 20]. Hinckley et al. compared users’ performances with either a standard mouse or a 6 degrees of freedom (dof) tracker, using an isomorphic mapping between the tracker and the virtual object rotations [11]. The tracker had either the shape of a small cube or a small sphere. They found that the shape did not matter. Users were found more efficient with the tracker than with mouse-based interactions, which demonstrated people’s ability to perform the integral control of the 3 dof of 3D rotations. Ware et al. tested various factors in the integral control of rotations [20]. They observed that having the hand in the same physical location as the virtual object was important, and they confirmed that the shape of the device was not. The isomorphic mapping used in both studies seems to provide the most natural interaction, in the sense that it is similar to the rotation of physical objects. However, it also suffers from the *limitations* of physical rotations: large rotations require large motor efforts. Poupyrev et al. studied non-isomorphic controls in order to introduce amplification in the transfer function [14]. They tested an amplification factor of 1.8 in a relative control, and observed that it was more efficient than the isomorphic control for rotations larger than 70°. Non-isomorphic amplification of rotation was also successfully used in the design of novel 6 dof devices [5]. Despite the demonstrated benefit of amplification, we chose not to use it with our spherical HPCD for two reasons. Firstly, any departure from an isomorphic control would break the rigid relation between the display and its virtual object; which may break the illusion of presence of the object. Secondly, the relative control used in previous studies breaks the “nulling compliance” [14]: rotations cannot be undone simply by moving the physical device back to its original orientation.

Object examination

Depth perception and 3D rotations were traditionally studied as two independent building blocks that can be used in a wide range of applications. In more recent works, depth perception and 3D rotations appeared as two *interdependent* components serving object examination.

Stavness et al. presented the first study on the capabilities of a cubic HPCD in an object examination task [16]. They tested a tree-tracing task in which the model could be rotated with various interactions. Two of the tested interactions used the cubic HPCD in an isomorphic control of the tree rotations. They offered either a direct interaction with the tree displayed in the HPCD, or an indirect interaction with the tree displayed on a standard monitor. Response time was higher in the direct interaction than in the indirect interaction, indicating a negative effect of displaying on the HPCD. In addition, the two interactions had higher response time than when controlling the tree rotation with a mouse. The best performances were achieved in a bimanual interaction using the HPCD for display and a mouse to control the tree rotations. While this provided encouraging results concerning the potential of HPCDs, we are focusing on HPCDs’ applications that are not constrained by the use of a mouse on a desktop. In addition, the pCubee prototype had technical limitations; which may have had strong effects in these results: only 5 of its faces were covered by LCD panels, it was connected to a computer using a thick wire, and it weighted 1.3 kg. This may have prevented fast and unconstrained rotations of the device. In addition, the LCD panels had borders that occluded a significant portion of the virtual content of the display, and stereo was not used. We extend this work by providing a high performance HPCD that removes most of the technical limitations of the pCubee prototype. In addition, we present the first comparison of the interaction with a HPCD and with physical objects.

Jansen et al. studied the exploration of 3D bar charts in information visualizations tasks involving the sorting and comparisons of cells of bar charts, and finding ranges of values [12]. They presented the first study that included physical objects in the comparison of interaction techniques. Physical 3D bar charts were found to offer superior performance compared to their virtual counterparts. This was mainly attributed the possibility of using touch on the bars, which provided strong cognitive aids in the data query tasks. But the superior “visual realism” of the physical bar charts was also suspected to favor this modality. We build on this work, but while Jansen et al. focused on *visualizations* of abstract data and data queries, we study the examination of *models* representing real objects. In addition, we contribute with a novel prototype that improves the sense of presence of the virtual objects.

A SPHERICAL HANDHELD PERSPECTIVE CORRECTED DISPLAY

Design Rationale

We designed our spherical display with the object examination experiment in mind. We aimed at high enough performances that participants would forget about the display and feel that they are actually holding a transparent display containing a rigidly attached 3D object. Using LCD or OLED panels as in pCubee [16] was quickly rejected because of the difficulty to cover the entire surface of the handheld display with no seams, and to create a lightweight *wireless* display. We opted for the projection of images on a passive object painted in white. External projection allows the use of a lightweight object that is very comfortable to move around. In addition, it offers great flexibility for choosing the shape of the display. We chose a spherical display rather than a cubic one to remove the discontinuities at the edges of a cube’s faces.

Implementation

We used a projector with 2560×1600 pixels at 120 Hz. By projecting small text, we tested that the image remained in focus even when moving the display ±15 cm along the projection axis from a focus position set manually. The projector was set above the participants’ head pointing downward with a roughly 60° angle from the horizontal, as illustrated in Figure 2. The custom-developed C++ software ran on an Intel Core i7@4 GHz desktop computer with an AMD Radeon R9 M295X graphic card. Tracking was performed at 240 Hz by a six-camera Optitrack^{TM} system. We used LCD shutter glasses for stereo display. The 120 Hz rate of the projector was split in 2×60 Hz for the two eyes, with independent rendering so that the most recent tracking data was always used.

The spherical display was made of a 14 cm ~0.1 kg transparent plastic sphere that we spray-painted in white. Ten passive markers were attached to the sphere and 4 to the shutter glasses for tracking purpose. The left and right eye positions were measured in a calibration step using an additional marker. We calibrated the projector using OpenCV *cv::calibrateCamera* function, so that we could illuminate any chosen point in the physical space. With the projector calibrated, we used a simple texture mapping approach to project the perspective corrected image of the 3D model on the spherical display. The strategy was to form, on a spherical mesh approximating the physical sphere, the perspective image that an eye would see on a simulated plane set between the eye and the sphere. This process is illustrated on Figure 3.

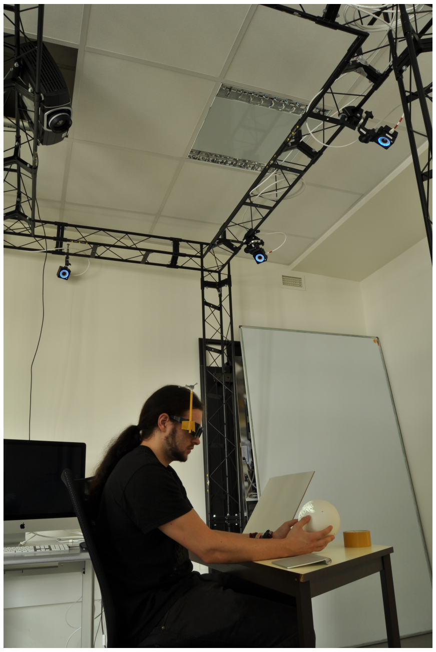

Figure 2. The projector and tracking setup.

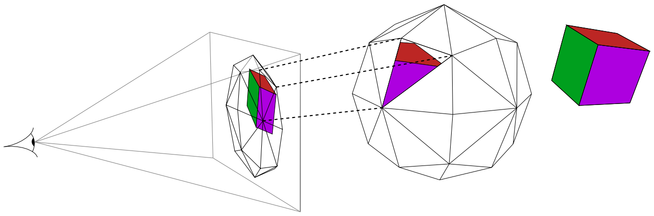

Figure 3. Perspective corrected display with texture mapping. The texture (base of the pyramid) is created from a virtual camera set at the eye’s position and looking at the center of the spherical display. For illustration purpose, a single triangle from a coarse spherical mesh is textured by back projecting its vertices in the camera’s projection plane. The projected scene (multicolor cube) is shown outside of the display, but the approach generalizes to any position, e.g. inside the display.

Performances

The use of 10 markers on the sphere and 6 cameras resulted in a *robust* tracking. During the study, we gave no indication on how to hold the sphere. Participants often held it with two hands, hiding many markers, but tracking failures were very sparse. *Accuracy* was measured at less than 2 mm by projecting a small sphere on a marker that we placed in various locations of the workspace. We empirically chose the parameter of the system to insure that the projected model had no discontinuities: the sphere approximation was made of 1536 triangles and was textured by a 1024×1024 image. The puzzle models were made of around 80 000 triangles. With these parameters, the total *rendering time* (updating the geometry and the texture coordinates, rasterizing) always took less than 4 ms. The end-to-end *latency* of the system was estimated with the simple predictive approach introduced by Cattan et al. [3]. We recorded eight latency estimations at 35 ms and two at 34 ms. The strong coherence of these independent measurements supports the accuracy of the estimation. This low latency still had a clearly perceptible effect when the spherical display was quickly shaken. This did not appear as problematic for our experiment as quick displacements were not required to perform the task: it was best executed with smooth rotations, allowing the participant's visual focus to follow a path in the puzzle. With this kind of manipulation the effect of latency was unnoticeable.

EXPERIMENTAL EVALUATION

In this experiment, we tested our hypothesis H that a Handheld Perspective Corrected Display (HPCD) improves users' efficiency in the examination of complex objects when compared to more common interaction techniques. In a main study, we compared the spherical display with two more common interactions that both relied on a planar display, and that used as input device either a touchpad or a tangible prop.

We were also interested in assessing how well the 3 digital interactions performed with respect to object examination in the physical world. In a secondary study, we included the interaction with a physical object in order to get a baseline performance.

Experimental setup

We used the spherical HPCD presented in the previous section. For the fixed planar display, in order to minimize differences in the image quality between experimental conditions, we used the same source of images (the projector) and the same type of display surface. We made a planar display from a wooden rectangle that we spray-painted with the same paint as the one used on the sphere. The projector's focus was optimized on the planar display.

We setup the experimental workspace on a small desk with the planar display on the left, a standard keyboard in front of it, and an Apple Magick Trackpad$^{\mathrm{TM}}$ to the right of the keyboard (all participant were right-handed). Space was saved to the right of the planar display and we asked participants to use the spherical display in this area: the focus of the projector was optimized to this distance, and this reduced the difference in pixel density between the planar and spherical display. We measured a median pixel pitch of 0.28 mm (91 dpi). This setup is shown in Figure 2.

Stereo and head coupling was used with all the virtual interaction techniques using the calibrated eye positions, as discussed previously. The inter-eyes distance was adapted to each participant by measuring it with a ruler. Stereo and head coupling were also naturally available in the physical condition.

A Puzzle-Solving Task

The perception of the structure of complex 3D objects has been measured in many studies using path-tracing tasks on complex graphs [1, 6, 16, 18, 19]. The graphs used often include hundreds of 1-pixel thin edges, which would represent a challenge to materialize as a physical object. We chose to adapt the path-tracing task to a puzzle-solving task where edges are replaced by thick curvilinear strings, or "spaghettis". Each puzzle is made of 6 intertwined spaghettis with a thickness of 5 mm, which can be 3D printed. One of the spaghettis has a black extremity. Participant must find the other extremity and tell its color. We randomly generated hundred models of puzzles and hand picked 10 of them for the experiment, plus some more for participants' training. We chose puzzles that offered a reasonable level of difficulty, as estimated subjectively by the authors.

Figure 4. The six physical puzzles used in the experiment. The one in the center has several "starts": it was used for training.

To get a baseline performance with physical objects, we created 6 physical puzzles by reproducing a subset of the virtual puzzles in a physical form, as illustrated on Figure 4. Five of the physical puzzles were used for performance measurement and the last one was used for participant's training. The spaghettis in the puzzle were 3D printed in white plastic. We built boxes to hold the spaghettis in place and to provide a convenient way for participants to hold and rotate the puzzles. The faces of the boxes were laser cut from 3 mm thick transparent acrylic. Using the puzzles' digital models, we also laser cut holes in two opposing faces of the cubes to hold the spaghettis. This allowed a faithful reproduction of the spatial configuration of the spaghettis. We observed that the edges of the boxes were the source of a bit of occlusion, but this did not seem to prevent an efficient inspection of the puzzles. We thought about building the puzzles inside transparent spheres instead of boxes to minimize the differences with the spherical display. However, we did not find a practical way to place and maintain the spaghettis inside a sphere while maintaining a faithful reproduction of their spatial configuration.

Participant used the keyboard to control the session. In the virtual conditions, participants pressed the spacebar to show a new puzzle, which was always presented in its start orientation: with the black spaghetti facing the user. In the physical condition, the operator gave the physical puzzle in its start orientation to the participants, but they were told not to look at it before they pressed the spacebar. After pressing the spacebar, time recording was started and participants visually followed the designated spaghetti until they found its color at its other end. They pressed the spacebar as soon as possible again to stop the time recording. At this point, the puzzle was hidden in the virtual conditions, or it was taken away from the participant in the physical condition. Participants then entered their answer by pressing one of the six keyboard's keys on which colored stickers had been attached, matching the 6 colors of the puzzles. A message was displayed to tell if the answer was correct or not. Then, participants either advanced to the next puzzle, or solved the same puzzle again depending if they provided a correct or incorrect answer.

Interaction techniques

The interaction technique was the main factor tested in our experiment: it was used to control the rotation of the puzzles. In the main study, we compared participants' performances in three conditions:

- In the `PAD` condition, the puzzles were rendered on the planar display. Rotations around the horizontal and vertical axes of the display were performed by sliding a finger vertically and horizontally on the touchpad, respectively. The control-display gain was set to 0.42 rad/cm. We chose it so that a single sliding gesture on the touchpad rotated the object by half a complete turn. Rotations around a perpendicular axis of the display could be performed by a rotation gesture of two fingers on the touchpad, the amplitude of the object rotation matching that of the fingers' rotation.
- In the `PROP` condition, the puzzles were also rendered on the planar display. Participant held the spherical display in their hands, but nothing was projected on it. It was only used as an input device to rotate the puzzle: the virtual puzzle was set to constantly match the orientation of the sphere with an isomorphic mapping and with an egocentric frame of reference [17]. Participants usually hold the sphere with two hands in front of them, while avoiding occluding their view of the puzzles.
- In the `HPCD` condition, the puzzles were rendered on the spherical display, with their center aligned with the display's center, and as if rigidly attached to it. Rotations of the display produced isomorphic rotations on the puzzles.

In the `PAD` and `PROP` conditions, the virtual puzzles remained centered on the planar screen. To keep the interaction as simple as possible, we decided not to add interactions for the control of translations: lateral translations were not considered useful for the task, and participants could lean on the display if they wanted to get a more detailed look on the puzzles. In the `HPCD` condition, all translations were possible to create the effect of the puzzle being contained in the display, however they were largely ignored during the experiment.

The `PAD` condition was selected as the "status quo". We preferred the use of a touchpad over a mouse because it supports the control of more degrees of freedom and multi-touch interaction is now widespread. However, although participants did not seem to have difficulties with the two fingers rotation during training, we observed that it was never used during the experiment.

The `PROP` condition was selected because it provided an intermediate design between the status quo and the spherical display. Comparing the `PROP` and `PAD` conditions allows isolating the benefits of the isomorphic rotations of the virtual object and the physical sphere. Comparing the `PROP` and the `HPCD` conditions isolates the benefits of having the hands physically in the same location as the virtual object.

In the secondary study, we included the `PHYS` *physical* condition were participants performed the task with the physical puzzles.

Experiment Design

20 unpaid subjects (4 female, median age 26.7 range [21..35], all right-handed) were recruited after screening for adequate stereovision using the Stereo Optical RANDDOT stereopsis test. All used computers and touch interaction (either direct or indirect) on a daily basis. After being introduced to the experiment's objectives and to the task, participant performed the experiment and then filled a post-experiment questionnaire.

We anticipated that improvement through training would occur both for the task and the interaction techniques. We thus split participants' sessions in 3 blocks. The first block was a short *training* block allowing participant to discover the task and the 4 interactions. Learning usually decreases exponentially, we expected the vertical part of the learning curve to occur in this training block. Participants were trained on 5 different training puzzles in the 3 digital conditions. In the physical condition, we used a single physical puzzle for the training, but we asked participant to solve successively 5 of its 6 spaghettis.

The two subsequent *measured* blocks were identical. Having two identical blocks allowed us to assess the remaining effect of training after the training block. In the main study, we used the same 10 virtual puzzles in both measured blocks and in all 3 conditions. In order to keep participant's sessions to a reasonable duration (under an hour), we interleaved the main study (comparing virtual conditions) and the secondary study (comparing to the interaction with a physical object): the data from 5 of the 10 puzzles was also used for the secondary study, in addition to the data from the 5 corresponding physical puzzles.

In order to prevent that participant remember the sequence of correct answer, the presentation order of the measured puzzles was randomized. In addition, the presentation order of conditions was balanced across participants. When switching to a new condition, participants were given a "warm-up run" on a non-measured puzzle to minimize the effect of the switch.

We recorded the `solving_time` as the time between the two presses on the spacebar *for the first attempt* at solving a puzzle. In case of error, we retained the solving time of the first answer, as subsequent answers benefitted from the previous examinations. We also recorded the `error_rate` as the number of incorrect answers over the total number of trials. Overall, we recorded for the main study 20 participants × 2 blocks × 3

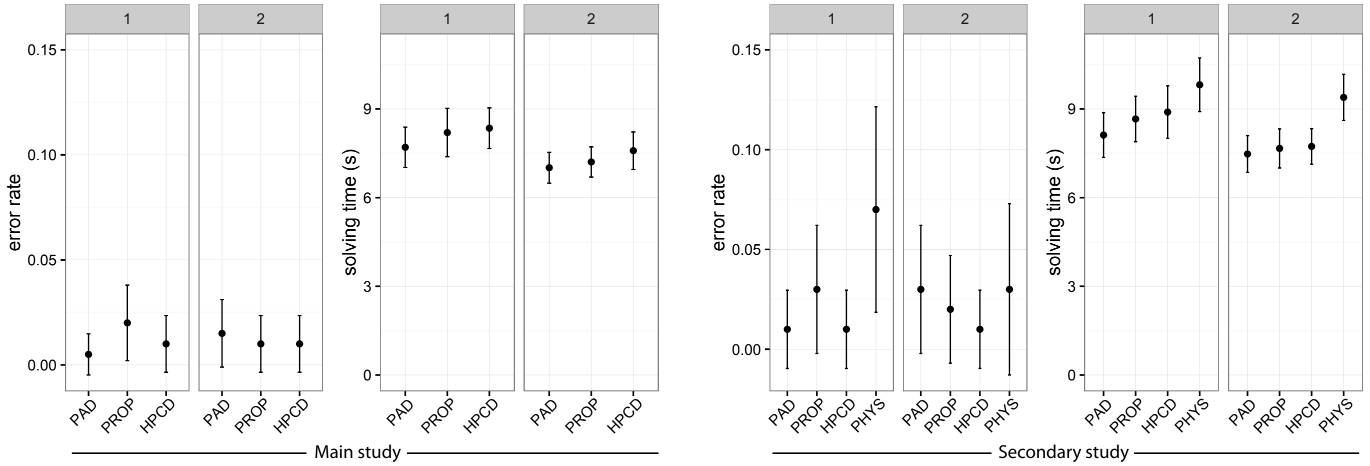

Figure 5. Results for the main (left) and secondary (right) studies: mean error_rate and mean solving_time per block and condition, with 95% Confidence Intervals (CI).

digital conditions × 10 trials = 1200 trials on *virtual* puzzles. For the secondary study, half of this data was used (5 virtual puzzles), plus an additional 20 participants × 2 blocks × 1 physical condition × 5 trials = 200 trials on *physical* puzzles.

Results

We split the data in 2 sets corresponding to the 2 studies. In the *digital* dataset, we only kept the data for the 3 digital conditions. In the *physical* dataset, we kept the data for all conditions but only for the 5 puzzles that had a physical embodiment.

Removing Homing Time

During the experiment, we realized that the use of the keyboard to start and stop the trials was better suited for the `PAD` condition than for the other ones: participants, all right-handed, kept their left hand ready to press the spacebar. When they found the puzzle's answer, they could immediately press the spacebar with no need to move their visual focus from the display. This was not the case in the other conditions, where the left hand was used with the right hand to manipulate the spherical display or the physical puzzles. Hence, pressing the spacebar required participants to first switch their visual focus to the keyboard and then to move their hand to the keyboard. This could take as long as 0.75 s. While this may be seen as a natural benefit of the standard desktop layout, we were interested in the more fundamental 3D examination time. Furthermore, HPCD design may include better-suited triggers such as taping a finger on the display. We thus decided to measure the homing time by post-processing the head tracking data and to remove it from `solving_time`.

We processed the head tracking data in the last few seconds before the press on the spacebar. The change of visual focus appeared clearly as a large spike in the head rotation and translation just before the key press. We used the beginning of this spike as the beginning of the homing. We used an automatic detection algorithm and then we manually checked all trials to correct a few miss-detections. Homing time was found very consistent per users and device, with an average across users at 0.250 s, 0.443 s, and 0.667 s for `PAD`, `PROP` and `HPCD`, respectively. There was thus a homing time also in `PAD`, which was due to participants beginning to orient their gaze towards the answer's key on the keyboard before stopping the trial with the spacebar. Head tracking was not used in the *physical* condition in the secondary study, but we observed that the homing behavior was very similar to the one in `HPCD`. We thus averaged the homing time in `HPCD` for each combination of user and block and used this average as an approximation of the homing time for `PHYS`.

Comparison of Digital Interactions

`error_rate` was low in all 3 conditions ($\leqslant 2\,\%$), it is represented on Figure 5, first graph. A repeated measure ANOVA did not show an effect of the interaction technique ($F(2,38) = 0.432, p = 0.652$), nor of the block (same number of errors, $p = 1.0$), or their interaction ($F(2,38) = 1.096, p = 0.344$). We thus concentrate our analysis on `solving_time`.

`solving_time` is represented on Figure 5, second graph. A repeated measure ANOVA revealed an effect of the block ($F(1,19) = 19.969, p < 0.001$), and an effect of the interaction technique ($F(2,38) = 5.95, p < 0.01$), but no interaction between the two ($F(2,38) = 0.423, p = 0.658$). Participants improved their performances between the block 1 and 2 by 9.0 %, 12 %, and 9.1 % in `PAD`, `PROP`, and `HPCD`, respectively. All improvements were significant, with $t(19) = 2.72, p < 0.05$ for `PAD` and $t(19) > 3.04, p < 0.01$ for `PROP` and `HPCD`. We thus focus our analysis on the second block, which better represents performances once participants became familiar with the task and the interaction. In block 2, pairwise t.tests with Holm correction revealed a single significant difference between `PAD` and `HPCD` ($p < 0.05$), with a solving time increase of 8.2 % from `PAD` to `HPCD`.

Comparison with Physical Objects

`error_rate` is represented on Figure 5, third graph. A repeated measure ANOVA did not find an effect of the interaction technique ($F(3,57) = 2.54, p = 0.066$), or an effect of the block ($F(1,19) = 0.234, p = 0.63$), or an interaction between the two ($F(3,57) = 1.20, p = 0.32$). Here again, we thus focus our analysis on `solving_time`.

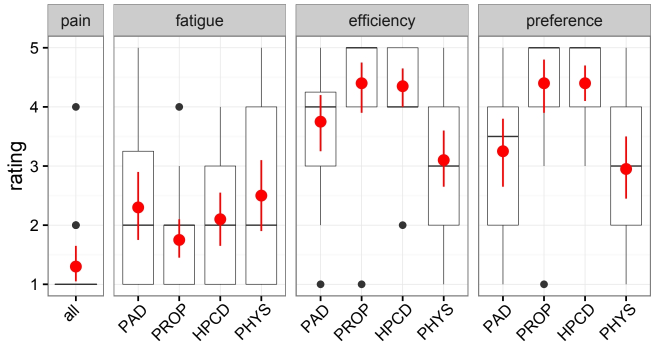

Figure 6. Participants' answers, from 1 "totally disagree" to 5 "totally agree". The exact questions are in the text. Boxplots are shown in black, means and 95% CI in red.

`solving_time` is represented on Figure 5, fourth graph. A repeated measure ANOVA revealed an effect of the interaction technique ($F(3,57) = 16.0, p < 0.001$), an effect of the block ($F(1,19) = 21.9, p < 0.001$), but no interaction between the two ($F(3,57) = 1.11, p = 0.352$). The performance improvement between the 2 blocks was significant for `PROP` and `HPCD` ($t(19) > 2.97, p < 0.01$), close to significant for `PAD` ($t(19) = 2.00, p = 0.060$), and not significant for `PHYS` ($t(19) = 1.36, p = 0.19$). The average `solving_time` reduced from block 1 to block 2 by 7.9 %, 11 %, 13 %, and 4.3 % in `PAD`, `PROP`, `HPCD`, and `PHYS`, respectively. Here again, we focus our analysis on the block 2 where participant were the most trained. In block 2, a pairwise t.tests with Holm correction revealed significant differences between `PHYS` and all three digital conditions ($p < 0.001$), but no significant differences between the digital conditions. Compared to `PHYS`, `solving_time` was improved by 20 %, 18 %, and 18 % in `PAD`, `PROP`, and `HPCD`, respectively.

Post-Experiment Questionnaire

In the questionnaire, we presented four sentences to participants and asked them to rate each with a number in the range 1 ("I totally disagree") to 5 ("I totally agree"). The first sentence addressed pain ("The experiment gave me pain in the eye, headache, or nausea.") and applied to the whole experiment. The three other sentences applied to each interaction independently, hence participants answered four numbers per sentence. The three questions addressed fatigue ("This interaction generate muscular fatigue."), efficiency ("This interaction allowed me to solve the puzzle efficiently."), and preference ("I like to use this technique."). Results of the questionnaire are summarized in Figure 6.

DISCUSSION

Superiority of the digital interactions over the physical interaction

We designed the secondary experiment as a way to assess how the three digital interactions performed with respect to what we thought would be the "gold standard": interacting with physical puzzles. Compared to their digital counterparts, physical puzzles offered a number of obvious advantages: infinite resolution, zero latency, perfect position accuracy for the head coupling and the eyes' stereo projections, and perfect matching of the eyes' accommodation and convergence. Results show that users' performances with the three digital interactions surpassed the performances with the physical puzzles. One limit of the experiment design is that participants performed more tasks in the digital conditions (27 runs per condition including training and warm-up) than in the physical condition (17 runs); which favored training in the digital condition. It should be noted however that people are highly trained at manipulating physical objects from their everyday life, and that little progress should be expected in the physical condition. This is coherent with the small (non significant) solving time decrease between the two blocks in the physical condition.

Post experiment comments from participants provide clues about factors that may have contributed to this result:

- 10 of the 20 participants reported that the occlusion created by the edges of the transparent boxes were perturbing. This did not prevent the solving of the puzzles (the average solving time in `PHYS` in only 18-20 % lower than in the digital conditions), but this may have slow down participants.
- 9 participants reported that the "contrast was not as good" on the physical puzzles as on the virtual ones. The natural lighting of the physical puzzles created inhomogeneous shading, with some orientation resulting in a faint visual separation of the spaghettis. Digital puzzles were rendered with standard artificial lighting (e.g. no radiosity), which was not photorealistic but always offered a clear separation between spaghettis.
- 2 participants reported that specular reflections could occur on the boxes' sides, and perturbed the search.

Other embodiments of the puzzles may have produce different results. For example, the occlusion from the boxes' edges would not have occurred had we managed to enclose the puzzles in transparent spheres, and simpler objects may not need any enclosure at all (e.g. [12]). Hence, one must be cautious about the generality of this result. However, the study revealed that the flaws of our virtual objects where outweighed by the flaws of the physical objects. Moreover, these physical flaws don't have trivial solutions: the spherical enclosure, avoiding specular reflections, and achieving a good shading under any orientation, all represent difficult challenges. From a designer's point of view, the study reveals an unexpected situation where the quest is no longer to improve a faithful virtual illusion of the physical world, but rather to compensate difficulties in the physical interaction. In other words, for close range visual examination, this study reveals that virtual 3D illusions have reached a level of quality that allows them to rival physical objects, and maybe surpass them.

Although for a different kind of tasks, Jansen et al. observed opposite results. They estimated that a better "visual realism" of their physical objects might have been a factor in their superiority over the digital counterparts [12]. Our study suggests that "realism" may not be the most important objective in virtual objects: because the shading of our virtual puzzles was *not* photorealistic, they offered better visual separation than the physical objects. As for the quality of the digital object, head coupling was not used in Jansen et al.'s study, and they estimated the latency of the isomorphic control at 100-150 ms vs. 35 ms in our system. As a result, the virtual objects in our study had better compliance to the physical world in term of structure from motion and in term of motor control, which may have contributed to the opposite result.

Non Superiority of the spherical HPCD

Contrary to our hypothesis, the performance with `HPCD` was *not* found superior to the performance with `PAD`: `HPCD` yielded 8.2 % *lower* performance than `PAD`. This was despite the fact that the interactions with the sphere (both in `HPCD` and `PROP`) were appreciated by participants: six participants expressed that rotation control with the sphere was "intuitive" and "efficient". Four participants expressed that the sphere permitted more accurate rotations than the touchpad. The questionnaire shows that participants had a tendency to find the interactions with the sphere the most efficient and the preferred ones, as illustrated on Figure 6.

`HPCD` having lower performance than `PAD` may appear as contradicting previous results showing the superiority of a tangible sphere [11] and of having the hand physically in the same location as the virtual object [20]. We attribute the different result to differences in the requirements of the experimental tasks. Previous studies on *3D rotations* used orientation-matching tasks; which involved the ability to anticipate the effect of the physical action on the device (i.e. the mapping between the device and the virtual object), and the ability to efficiently execute precise rotations. Both requirements would probably benefit from the intuitive isomorphic control offered by the sphere. However, in our experimental task, rotations are only a means to achieve the visual tracing of a spaghetti. Frequent changes of viewpoint are required in order to disambiguate the crossings of spaghettis, but this does not require a *specific* rotation: any rotation towards a coarse direction aimed at by the user provides sufficient *structure from motion* for the disambiguation. Hence, rotations in our experiment needed to be neither specific nor accurate. The experiment led us to understand that intuitive and accurate rotations were not important factors in the performance of the object examination task. This is coherent with findings from Jansen et al. who observed that an isomorphic prop condition was not superior to mouse interaction in *data query* tasks [12]. Our study indicates that this result extends to more general object examination tasks.

These loose requirements on rotations may also explain why rotations around the normal axis of the display were never used with `PAD`: the task could be achieved without them and the use of two fingers required additional motor efforts. On the contrary, being able to *quickly* rotate the puzzles may have been beneficial. We studied if the control-display gain offered by `PAD` contributed to faster puzzle rotations in the experiment. We post-processed event recordings from the experiments: we computed the mean and maximum rotation speed per device and participants. While the means were similar, the maximum speed averaged across participants was $6.88\,\mathrm{rad\,s^{-1}}$, $4.65\,\mathrm{rad\,s^{-1}}$, and $4.21\,\mathrm{rad\,s^{-1}}$ for `PAD`, `PROP`, and `HPCD`, respectively (all pairwise comparison significant with $p < 0.01$). The maximum speed was thus 48 % and 63 % higher in `PAD` than in `PROP` and `HPCD`, respectively. Rotation profiles with `PAD` revealed a pattern of quick changes interleaved with pauses, while changes were smoother with `PROP` and `HPCD`. The high performance of `PAD` in our experiment indicates that this pattern is well suited for object examination.

The visual quality of the virtual puzzles may also have had a detrimental effect on `HPCD` efficiency: five participants expressed that the image was blurry when looking at the periphery of the sphere, which reduced the extent of the usable area for working with the puzzles. This was a limitation of our prototype due to the use of projection. In addition, four participants reported an annoyance due to self-shadowing, which prevented the sphere to be held too close from the body.

The Potential of HPCD

Although the `HPCD` did not improve performance on the task of object examination, participants' comments and ratings indicates that it has strong potential to perform other tasks that require an intuitive control of precise and accurate rotations. The study of a 6 dof docking task [5] appears as a promising next step in the assessment of HPCDs benefits. In addition, the experimental task focused on the rotations of a display that remained mostly in the same location. But moving the display around appeared to be quick and intuitive. This points to a different context where the virtual scene is not limited to the boundary of the display: HPCDs should be tested as mobile tangible *volumes*. This would extend the concept of spatially aware 2D displays for the exploration of situated information space [4] and the concept of tangible planar windows for the exploration of 3D spaces [15]. Finally, six participants commented that "the illusion of the object inside the sphere was perfect". This points to a formal evaluation of the illusion of presence of the virtual object, and to promising applications in teaching and museum contexts.

CONCLUSION

We introduced a novel approach to the implementation of handheld perspective corrected displays (HPCDs) using external projection. It allowed the creation of a very maneuverable spherical HPCD that created a strong illusion of a rigidly attached virtual object inside the display. In spite of allowing intuitive, fast and accurate rotations of the virtual object, the display did not improve participant's performances in a complex examination task. The study indicates that coarse rotations are sufficient for object examination, and it points to other tasks that should benefit from HPCDs such as 6-dof docking or scene exploration through a tangible volume.

ACKNOWLEDGEMENTS

This work was supported by the ISAR project ANR-14-CE24-0013.

REFERENCES

1. Kevin W. Arthur, Kellogg S. Booth, and Colin Ware. 1993. Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems* 11, 3 (1993), 239–265.

2. Ragnar Bade, Felix Ritter, and Bernhard Preim. 2005. Usability comparison of mouse-based interaction techniques for predictable 3d rotation, In Proceedings of the 5th international conference on Smart Graphics. *SG* (2005), 138–150. `DOI:` http://dx.doi.org/10.1007/11536482_12
3. Elie Cattan, Amélie Rochet-Capellan, and François Bérard. 2015. A Predictive Approach for an End-to-End Touch-Latency Measurement, In ACM Interactive Tabletops and Surfaces. *ITS* (2015). `DOI:` http://dx.doi.org/10.1145/2817721.2817747
4. George W. Fitzmaurice. 1993. Situated Information Spaces and Spatially Aware Palmtop Computers. *Commun. ACM* 36, 7 (1993), 39–49. `DOI:` http://dx.doi.org/10.1145/159544.159566
5. Bernd Froehlich, Jan Hochstrate, Verena Skuk, and Anke Huckauf. 2006. The GlobeFish and the GlobeMouse: two new six degree of freedom input devices for graphics applications, In ACM Conference on Human Factors in Computing Systems. *CHI* (2006), 191–199.
6. Tovi Grossman and Ravin Balakrishnan. 2006. An evaluation of depth perception on volumetric displays, In ACM Conference on Advanced Visual Interfaces. *AVI* (2006), 193–200.
7. Mark Hancock, Thomas ten Cate, and Sheelagh Carpendale. 2009. Sticky tools: full 6DOF force-based interaction for multi-touch tables, In ACM International Conference on Interactive Tabletops and Surfaces. *ITS* (2009), 133–140. `DOI:` http://dx.doi.org/10.1145/1731903.1731930
8. Karin L. Harman, G. Keith Humphrey, and Melvyn A. Goodale. 1999. Active manual control of object views facilitates visual recognition. *Current Biology* 9, 22 (2016/08/30 1999), 1315–1318. `DOI:` http://dx.doi.org/10.1016/S0960-9822(00)80053-6
9. Otmar Hilliges, David Kim, Shahram Izadi, Malte Weiss, and Andrew Wilson. 2012. HoloDesk: Direct 3D Interactions with a Situated See-through Display, In ACM Conference on Human Factors in Computing Systems. *CHI* (2012), 2421–2430. `DOI:` http://dx.doi.org/10.1145/2207676.2208405
10. Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. Passive real-world interface props for neurosurgical visualization, In ACM Conference on Human Factors in Computing Systems. *CHI* (1994), 452–458.
11. Ken Hinckley, Joe Tullio, Randy Pausch, Dennis Proffitt, and Neal Kassell. 1997. Usability analysis of 3D rotation techniques, In ACM Symposium on User Interface Software and Technology. *UIST* (1997), 1–10.
12. Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. 2013. Evaluating the Efficiency of Physical Visualizations, In ACM Conference on Human Factors in Computing Systems. *CHI* (2013), 2593–2602. `DOI:` http://dx.doi.org/10.1145/2470654.2481359
13. Frank Meijer and Rob H.J. Van der Lubbe. 2011. Active exploration improves perceptual sensitivity for virtual 3D objects in visual recognition tasks. *Vision Research* 51, 23–24 (2011), 2431 – 2439. `DOI:` http://dx.doi.org/10.1016/j.visres.2011.09.013
14. Ivan Poupyrev, Suzanne Weghorst, and Sidney Fels. 2000. Non-isomorphic 3D Rotational Techniques, In ACM Conference on Human Factors in Computing Systems. *CHI* (2000), 540–547. `DOI:` http://dx.doi.org/10.1145/332040.332497
15. Martin Spindler, Wolfgang Büschel, and Raimund Dachselt. 2012. Use your head: tangible windows for 3D information spaces in a tabletop environment, In ACM International Conference on Interactive Tabletops and Surfaces. *ITS* (2012), 245–254. `DOI:` http://dx.doi.org/10.1145/2396636.2396674
16. Ian Stavness, Billy Lam, and Sidney Fels. 2010. pCubee: A Perspective-corrected Handheld Cubic Display, In ACM Conference on Human Factors in Computing Systems. *CHI* (2010), 1381–1390. `DOI:` http://dx.doi.org/10.1145/1753326.1753535
17. Colin Ware and Roland Arsenault. 2004. Frames of Reference in Virtual Object Rotation, In ACM Symposium on Applied Perception in Graphics and Visualization. *APGV* (2004), 135–141. `DOI:` http://dx.doi.org/10.1145/1012551.1012576
18. Colin Ware and Glenn Franck. 1996. Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Trans. Graph.* 15, 2 (1996), 121–140.
19. Colin Ware and Peter Mitchell. 2005. Reevaluating Stereo and Motion Cues for Visualizing Graphs in Three Dimensions, In Proceedings of the 2Nd Symposium on Applied Perception in Graphics and Visualization. *APGV* (2005), 51–58. `DOI:` http://dx.doi.org/10.1145/1080402.1080411
20. Colin Ware and Jeff Rose. 1999. Rotating Virtual Objects with Real Handles. *ACM Trans. Comput.-Hum. Interact.* 6, 2 (June 1999), 162–180. `DOI:` http://dx.doi.org/10.1145/319091.319102