Perceptual User Interface Framework for Immersive Information Retrieval Environments

(An Experimental Framework for Testing and Rapid Iteration)

Markos Mentzelopoulos [1], Jeffrey Ferguson [1], Aristidis Protopsaltis [2]

[1] University of Westminster, United Kingdom
[2] Friedrich-Alexander-Universität Erlangen-Nuremberg, Germany

Abstract—**The use of perceptual inputs is an emerging area within HCI that suggests a developing Perceptual User Interface (PUI) that may prove advantageous for those involved in mobile serious games and immersive social network environments. Since there are a large variety of input devices, software platforms, possible interactions, and myriad ways to combine all of the above elements in pursuit of a PUI, we propose in this paper a basic experimental framework that will be able to standardize study of the wide range of interactive applications for testing efficacy in learning or information retrieval and also suggest improvements to emerging PUIs by enabling quick iteration. This rapid iteration will start to define a targeted range of interactions that will be intuitive and comfortable as perceptual inputs, and enhance learning and information retention in comparison to traditional GUI systems. The work focuses on the planning of the technical development of two scenarios, and the first steps in developing a framework to evaluate these and other PUIs for efficacy and pedagogy.**

Index Terms—**perceptual; Perceptual User Interface (PUI); Human Computer Interaction (HCI); serious games; Graphical User Interface (GUI)**

I. INTRODUCTION

The computer hardware and software development vector indicates movement away from a traditional windows, icons, menus and pointer (WIMP) and desktop paradigm based on 2D content and (traditional GUI) interaction. Input and manipulation of this variety is mature at present, with the mouse and keyboard as 1D or 2D non-perceptual interfaces that have been widely accepted for some time by the general public and developers. The extension of our interactive experience to mobile computing has brought touch screen and 2D gesture language that the majority of users are comfortable with and believe enhances their experience. Further developments with 3D and stereo imaging open the possibilities of a richer immersive environment more alike our perception of the physical world. 3D graphics capabilities are now integrated into most PCs and mobile devices, along with the webGL framework for html5, which can start to standardize 3D objects and interaction on the web platform. When full 3d representational capabilities begin to exist seamlessly across all platforms and interfaces, rather than continue with the current 2D abstracted imagery, there will be a low barrier to creating a far richer interactive experience. The first instances of this are currently being explored using a new range of commercial immersion sys-

tems such as Project Morpheus by Sony, Hololens by Microsoft, and Oculus Rift by Facebook. To enable interaction with these widely used immersive environments, the mouse and keyboard with associated 2D GUI are adequate starting points, but developers look toward an advancement for interaction devices on par with that of the software and hardware platforms.

Perceptual interfaces such as Kinect and LEAP motion (gesture recognition), other hardware sensors and inputs such as Razer Hydra (motion capture with triggers) and the Oculus Touch (hybrid) combine tracking of the hands and gesture recognition or trigger activation. These take interaction outside of the GUI's two dimensions and offer some more complex mapping-to-control in a 3D environment away from an anchored desktop hardware setup.

Interactive 3D engines such as Unity and Unreal make development of immersive environments and integration of commercial controllers possible without a great deal of bespoke coding. However, rather than being developed with the specific purpose of advancing the educational experience, this technology is being advanced with or without that input. Therefore, it is important to try to promote a dialogue around pedagogical method with analysis of improvements in education that might be gained from specific technical scenarios.

Starting a dialogue here with two scenarios developed using current accepted commercial hardware and software keeps the analysis and discussion at a practical level with these developers. The two technical scenarios are an immersive social network and a learning environment suitable for a museum kiosk. Both set out to explore a comfortable Perceptual User Interface that is a natural progression from the 2D GUI [1]. Key elements that are included in both of our technical scenarios are; the use of a game quality 3d engine (with or without stereo or HMD use), gestural or motion capture input from devices available to the general public, and finally, content familiar to users of social media, the internet, or museum kiosks.

The idea is to not overwhelm the senses or tire the user, but instead find optimal, efficient, and natural interfaces for immersive environments. Some of this can come from past lessons in digital puppeteering, HCI taxonomy, and cognitive and perceptual motor interaction. Due to the range of perceptual inputs available commercially, the large amount of software platforms and the variety of interaction or learning software that can be coded using both together, there is potential for very many "recipes" for interactive applications using perceptual interfaces.

Since we are at such an early developmental stage, it will benefit us as researchers of education applications to be able to quickly quantify the efficacy of each new wrinkle of learning software or interface. Comparisons between the ranges of educational software output will enable rapid iteration of the more successful software and hardware interfaces, and may suggest new combinations.

As an end goal, a new, more efficient PUI will be ideal for the new richer immersive environments. Gesture recognition, voice commands, and eye-tracking all present themselves as lower level inputs that do not tire out the user, but offer a good cognitive control-to-task fit. It would be ideal to find optimal setups and guidelines for these interactions.

II. TWO CASES

In order to set out an experimental framework to study and compare learning and information retrieval applications using perceptual interfaces, we present two test cases that are largely indicative of two ends of the software, hardware, and usage spectrum. The first is a museum learning application that uses a depth camera for gesture interaction and information retrieval. The second is a conceptual interface for an immersive social media application using motion capture devices typically used for gaming.

A. Interactive Solar System

The first scenario is an implementation of a typical information presentation/retrieval system for learning software or a museum kiosk. This tracks the users hand positions and offers navigation of both a 2D menu and a 3D scene using gesture recognition. The 3D scene is limited in navigation to aid gestural control, and depth camera sensing of grasps controls selection. Display is by a 2D screen in front of the standing user.

Specifications for the system are listed below:

- Kinect 2, Unity 3D 5.0.0fe, Visual Studio 2013
- Gestural interface
- Interactive navigation of 2D menu by hand tracking
- Selection by grasping gesture recognition
- Movement between fixed 3D positions in 3D scene
- Manipulation of 3D objects by grasping for selection and moving hand for rotation.

The first technical scenario was developed using the Unity Engine with Kinect 2 interaction scripted in C#. Kinect 2 offers a significant development advance over the original Kinect in terms of resolution and integration with a range of software. The Kinect 2 was chosen for its ease of integration with the Unity 3D engine without the need of further 3rd party plugins, and the fact that hand gestures could be easily captured and accessed. When the program runs, the user is first presented with a menu screen which is effectively a 2D menu (Fig.1). Both of the user's hand positions are tracked and translated to 2D icon positions on the screen. If a hand icon intersects with a button or object in the 3D space, the hand icon is highlighted. Hand grasps are recognized as gestures, trigger a grasp icon, and serve as button presses. By pressing one of the three buttons on the menu screen, the user can view a tutorial screen, exit the program, or start a 3D solar system simulation (Fig.2).
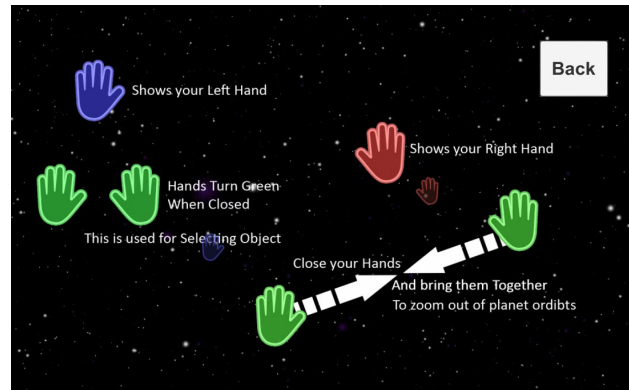

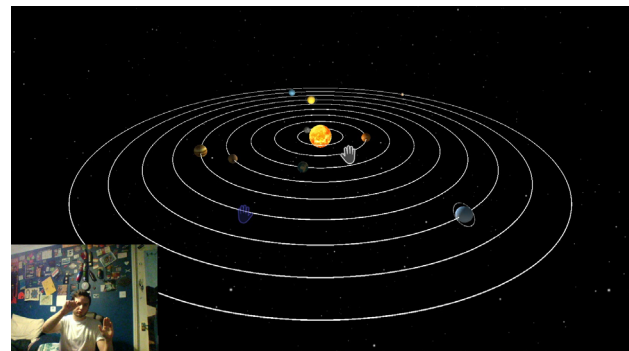Figure 1.   Interactive Solar System- Main Menu.

Figure 2.   Interactive Solar System- Simulation.

The 3D solar system simulation contains a camera at a vantage point for viewing the whole system. 3D planets rotate in orbits around the sun with accurate timings (Fig.3).

By selecting and grasping an orbiting planet, the user finds that the camera zooms in to a close view of the planet with an information panel (Fig.4). The user can then interact further by grasping and rotating the planet for a full view, or exiting back to the solar system view.

This scenario represents the current state of interaction with depth based perceptual devices such as the Kinect 2. The user may experience some variance of interaction depending on the accuracy with which the Kinect 2 picks them up and continues to track them. Issues can include the user's scale and position in relation to the capture device, and any environmental conditions. This gesture interaction method can be tested for learning retention and satisfaction against the same system with a traditional mouse or touchpad input GUI.

From looking at the parameters of the system, interaction is largely natural. The menu systems are successful, with the planet selection and rotation intuitive once the grasping concept is learned. Grasping is not a completely natural fit for the user, but is within acceptable Microsoft Human Interface Guidelines [12].

Elements that would further improve the interaction would be more reliable finger and hand tracking. This could bring the expansive gestures in. The sometimes expansive and tiring gestures are not ideal, and could be reduced in line with the digital marionette concept [5]. Concepts explored were around a 2D/3D information retrieval and learning system with gesture controls at an interactive software kiosk.
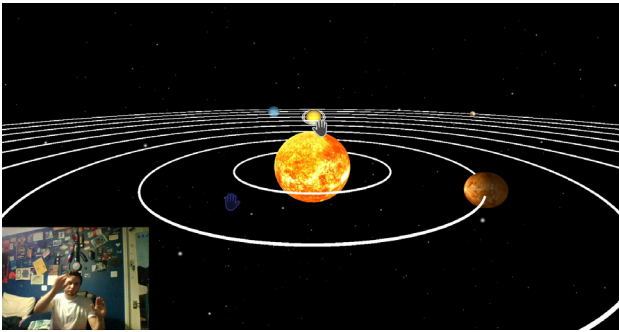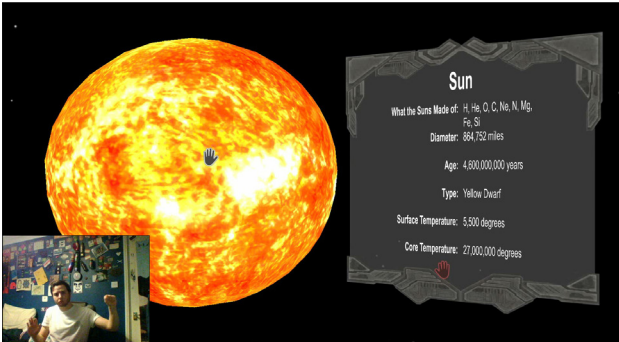
Figure 3.   Interactive Solar System- Zoom

Figure 5.   Interactive Facebook VR – Razer Hydra Callibration

Figure 4.   Interactive Solar System- Planet Panel Information

Figure 6.   Interactive Facebook VR – Hand Animation based on Razer Hydra motion feedback.

From the initial use, a study or experiment could be suggested on the comfort and match of the gestures to the task at hand, with further feedback from users regarding comfort, intuitiveness, understanding of expectations, and suggestions. Efficacy of the system versus traditional a GUI system could be compared.

B. Interactive Facebook VR

The second scenario is a conceptual implementation of a 3D immersive interface for Facebook. This also tracks the users hand positions and offers navigation of a full 3D interface, but using motion capture rather than gesture recognition. There are hardware triggers rather than perceptual gestures for grasping selection, and the display can be via an Oculus Rift as well as a 2D screen.

- Specifications for the system are listed below:
- Razer Hydra, Oculus Rift, Unreal Engine 4.7, Visual Studio 2013, Facebook simulator software Razer Hydra, Oculus Rift, Unreal Engine 4.7, Visual Studio 2013, Facebook simulator software [13]
- Motion Capture for hand position
- Selection by hardware trigger for grasping
- Navigation of 3D scene - unlimited X, Y and Z translation, bounded and guided by framework
- Pop up 2D menu system for calibration, exit, restart
- Stereo HMD for optional display

The second technical scenario was developed using Unreal Engine with Razer Hydra interaction scripted in C++. The Razer Hydra offers a more stable 3D tracking system than the Kinect 2, and hardware triggers make the system less gestural, but also more dependable. Rather than suffering from dropout as in the Kinect, the only issues with the Hydra are calibration and drift, both more easily correctable. The display is by Oculus Rift or computer screen.

The added reliability and integrated triggers of the hydra allows a fully 3D interface with unlimited navigation. This makes a good counterpoint to the Kinect experience for testing, as the Kinect would require a more cognitively complex control system to achieve the same results, yet not be as reliable. Each system, perceptual and mocap, is configured to its optimal ability. The framework scenery and post spawning behavior keeps the user at the center of the interaction despite the full freedom of movement.

To begin, the user picks up the Hydra controllers. When the program runs, the user is presented with a pair of 3D hands and the 3D framework scenery, stretching to infinity (Fig.5). Both of the user's hand positions and rotations are tracked and translated to 3D animated hand models on the screen. Hand grasps are animated by finger triggers on the front of the devices, and work conceptually well with no lag or error. The rotation and position of hands into and out of the screen provide satisfying feedback. The position of the hands into the screen also controls a depth fog effect, enhancing the user's investment with the interface as interactive (Fig.6).

The user controls his position by the left thumbstick, which moves him forward and back in space. In the version using the Oculus Rift display, user rotation is handled by the Rift sensor, and the right thumbstick controls the up and down motion of the user. In the screen display version, the right thumbstick controls the user rotation.

By pressing the start buttons, the user has the option to change settings such as hand calibration and Oculus Rift settings. There are times where the calibration drifts and needs to be brought back in line, or a user can calibrate his input range to one that is most comfortable for him.

The final element of the engine is the post generator. Images are simulated by a separate program, and mapped onto polygons in the Unreal Engine that are pickable ob-

jects (Fig.7). If the user presses the trigger to activate the grasp animation when in range of a post with either hand, the post is grabbed and held as long as the trigger is depressed. Releasing the trigger releases the post. Posts are generated as the user begins and continues to navigate. If the user is still, no posts are generated and current posts die out so that the environment does not become saturated.

This scenario represents the current state of interaction with motion capture game controllers in a 3D environment. Calibration is generally straightforward, but can be altered accidentally by the user so as to be unusable. Otherwise, with good calibration the system is satisfying in its responsiveness and accuracy.

Users generally find the hand movement and navigation satisfying, but have to work to come to grips with the grasping of actual posts in the environment as well as navigating to get in range. There is some spatial disconnect with collision that could be addressed, but sometimes calibration and practice leads to better handling.

The system can also be tested against a mouse and keyboard interface to determine if a more immersive motion tracking control system is advantageous. This can be compared to the gestural vs GUI comparison for the Kinect 2 in system 1.

Concepts explored were around a 3D immersive environment navigated by motion capture gaming device with social media content selectable by hardware triggers.

From the initial use, a study is planned to test the comfort and match of the movements to the task at hand, with further feedback from users regarding comfort, intuitiveness, understanding of expectations, and suggestions. Efficacy of the system versus traditional a GUI system could be compared.

The system can be tested both with and without the Oculus Rift for a comparison of user satisfaction of the interface and interaction with and without immersive stereo view. This is appropriate for the interaction and also the content, as Facebook has acquired oculus rift, and is likely testing immersive environments along a similar line to those of Microsoft's Hololens and project Morpheus by Sony. The control of the POV by either the Oculus Rift sensor or the right thumbstick would be ideal.

III. INVESTIGATION

To investigate educational opportunities of immersive software, there should be some structure by which to make comparisons, and therefore improvements to the rapidly expanding combinations of hardware and software interactions.

Initially, we can just compare the newer PUIs to the older GUIs. There is some perception that the newer PUIs would be more appealing to users. In fact, there is study that shows that users perceive that they accomplish tasks better when using the more familiar and physical devices for GUI [7].

Users considered 3D input more tiring, and the mouse easier. They also thought they did better on the mouse, even when they didn't [7]. While the gestural device is better for more immersive UIs due to the greater variety of control it presents, the mapping of its capabilities to the task are key, and these must be considered.

Figure 7.   Interactive Facebook VR – Post Interactivity

To paraphrase Jacob and Sibert; a taxonomy, or descriptive framework for pragmatic selection of input devices assists in a formal study of incorporating input devices into interaction frameworks [7, 12]. Perceptual structure is the key to understanding performance of multi-dimensional input devices on multi-dimensional tasks. Therefore, perceptual structure must be part of any experimental framework, or taxonomy, for these devices [7].

Between perceptual devices, Traditional HCI study and the newer PUI study provides some concepts to develop categories for comparison. According to Chua [6], translation, coding and mapping are the key to HCI. Translation is the human interface between perception and action, or stimuli and responses. Coding of the user input to stimuli and mapping to responses is the workflow to get correct in order to have an interface that is cognitively suitable.

Turk [1] further says that "The ideal user interface is one that imposes little or no cognitive load on the user, so that the user's intent is communicated to the system without an explicit translation on the user's part into the application semantics and a mapping to the system interaction techniques."

As Sturman and Zeltzer confirm [3], coordination of many degrees of freedom (dof) increases the cognitive workload, but good or task control mapping for devices reduces the learning curve and increases efficiency.

Fitts Law [15], a predictive model of testing time to engage a target, is the traditional method to measure the efficacy of an interface for selection on a single object. With more complex interfaces, there is also a cognitive process to test selection from an array in the Hick-Hyman Law[15]. All this can inform our precise questions to determine efficacy between PUI and GUI for correct perceptual structure developments.

Also, Effective PUI comparisons will enable comparison between the Oculus Rift sensor for POV rotation, or the Razer Hydra thumbstick. Comparison of PUI vs GUI will capture the visualisation difference between the Oculus Rift and a traditional 2D screen.

Finally, between the two scenarios presented here, there can be a comparison via this experimental framework of the users' comfort with navigating a fully 3D environment vs. a structured 3D environment (2.5D) that is represented by the Solar system. They may also prefer the reliability of the Hydra, but how practical is its use? Our framework should be able to capture this.

IV. EXPERIMENT FRAMEWORK

The next step of this work is to investigate how the varieties of these immersive environments with new control systems will compare with a traditional GUI interface and other PUI interfaces. The framework we propose in order to facilitate study of a variety of PUI scenarios and make suggestions for further rapid iteration is as follows:

1. Comparison between new PUI to old GUI for efficacy:
   - Investigate learning with new control systems vs. old
   - *Non Perceptual Preference* [7, 9]
   - Information Retention
   - Natural, Intuitive, Adaptive, Unobtrusive [1]
   - *Fitts law test & Hick-Hyman Law* [6,15]
2. Investigate Perceptual Structure[7,9]:
   - Cognitive Load [1]
   - Conceptual Space Disconnects[1, 3]
   - *Control to Task mapping*[3]
3. Investigate Motor Coordination:
   - Perceptual Motor Interaction Coordination [5,6,7]
   - *Ideomotor, Common Coding Theory*
   - *Conflicts*[3]
   - *Tiring*[3]
4. Comparison between Device Perceptual Structures:
   - Taxonomy (task based) [7, 9, 12]
   - Device Efficacy Perceptions (task based) [7, 9, 12]
   - Fitts law test, Hick-Hyman Law (task based) [6,15]

With the information gained by using this experimental framework on the wide variety of PUI applications, we should be able to suggest refinements to software and hardware parameters for further iteration.

V. DESIGN OF EXPERIMENTS

PUIs are being developed as a natural succession to GUIs. The multi-dimensional nature of the emerging immersive environments and our increasingly mobile interaction with these seems to indicate we must transition to a new HCI in order to effectively utilize them.

However, as shown by Jacob and Sibert [7], there is a continuing perception by users that tasks, even those in an immersive environment, are more easily accomplished by traditional GUI devices. This persists even when the evidence points to the contrary. One aspect of this that is hard to refute is the fact that users generally find the use of PUI input devices as more tiring [7].

Testing interfaces for efficacy and information retention will assist in comparing and iterating development of PUIs. The goal is to develop a framework for comparison, especially in regards to information retention and pedagogy, which can run in parallel.

The first steps toward a framework offering effective comparison of differing PUI systems will start with a PUI vs. GUI comparison. This PUI to GUI comparison should consider the following:

1. Effective Task/Control mapping (Motor and Cognitive)

2. PUI or GUI Preference

3. Pedagogical Efficacy

PUI or GUI preference is fairly straightforward to capture with sentiment analysis. Likewise, pedagogical efficacy can be quantified by using the same software with various PUI and GUI setups. Finally, any tiring effects of PUIs is also captured by sentiment analysis.

The real issue for analysing GUIs, PUIs, and comparing PUIs to other PUIs is finding an effective procedure for capturing a full range of comparison data around task to control mapping.

Fitts Law and the Hick-Hyman Law seem ideal for efficacy of menu or object selection using various devices. This can contribute in some way to a study between PUI, GUI for basic motor and cognitive task/control mapping. A follow-up set of questions based on information retention could start to capture the pedagogical efficacy range between systems. This would of course require naïve subjects for each separate input device, or a variation of task information within the system.

A deeper conceptual analysis is supplied by the Jacob and Sibert experiment directly comparing the conceptual frameworks of a mouse based GUI and a motion tracking based PUI. For their taxonomy, Jacob and Sibert expand Garner [7] for the 3D input from a magnetic tracking system.

They investigate the differing perceptual structures of multi-dimensional spaces, and how different devices engage with these structures. Their hypothesis is that "the structure of the perceptual space of the interaction task should mirror that of the control space of the input device."[7].

By expanding the Garner theories, they identify attributes of objects in multi-dimensional spaces. This defines their perceptual space. The relationship between attributes can be defined as either integral or separate, depending on how well the components remain identifiable. Those that perceptually blend together are integral. Those that do not are separate.

The motion tracking PUI in Jacob and Sibert [7] is the same technology to the Razer Hydra used for our virtual social network. It also has similarities to the Kinect 2, though these are based on the fact that both the Hydra and the Kinect 2 have perceptually integrated dimensions. Actual selection interaction of the Hydra is more akin to a gamepad, and that of the Kinect is purely gestural. That of the Hydra is closest to the gestures described by Jacob and Sibert [7] in their concluding zoom and pan task/control application example. This is fully integral in 3 dimensions. That of the Kinect is similar in hand tracking, but in 2 dimensions for our software example.

For a deeper experimental framework what is needed initially is an overview of use sentiment between differing PUIs. With the two systems we have, a direct comparison of two PUIs with integrated dimensions that are similar but with variation of composition will start to point the way to framework parameters.

For the evaluation, two perceptual user interfaces were compared. Both are Euclidean in nature, meaning the movement in 2 or 3 dimensions is integral rather than separate or stepped. The main difference in control input is that the Razer Hydra is a motion tracked device with selection interaction in 3 dimensions via button press in
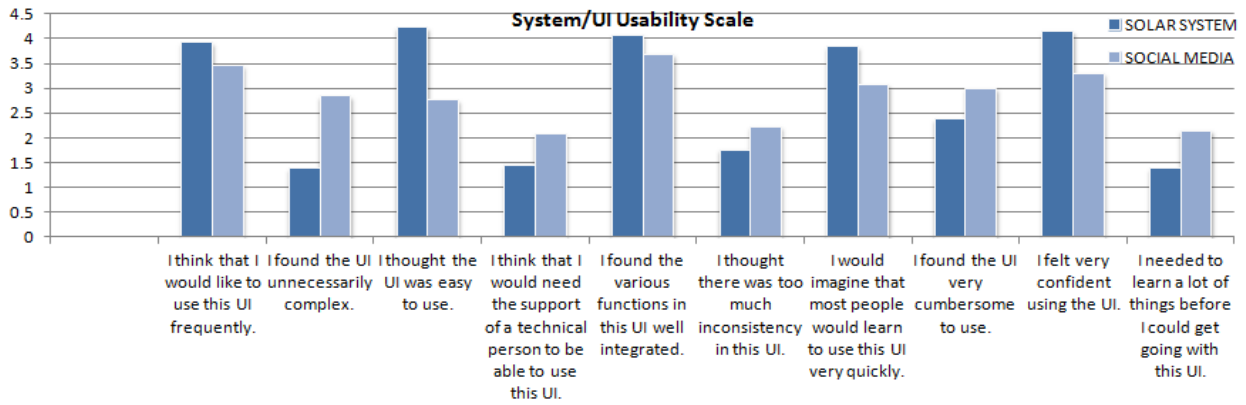
Figure 8.   UI Usability Scale

our software, and the Kinect 2 is a motion tracked device with selection interaction in 2 dimensions with gestural selections. Therefore in the sentiment analysis, the main differing component is the number of dimensions being navigated. Navigation for both systems is Euclidian.

So, apart from a general sentiment analysis to begin developing the framework for PUI comparison, the specific difference in this analysis is the extra Z dimension in the Social Media simulation.

VI.   EVALUATION

A.   Evaluation procedure and apparatus

The pilot evaluation of the UIs was carried out with 13 subjects (11 undergraduate, one post graduate and one researcher) with experience in the hardware used. The study took place at the University of Westminster, London premises and each participant was tested individually. Each session lasted for approximately 20 to 30 minutes. The participants had to use each system for 10 minutes and then answer a short questionnaire. The questionnaire consisted of 20 questions in total. All the questions were multiple choices on a Likert scale of one to five (one being the least favourable answer and the five the most favourable answer). The evaluation focused on usability issues, system capabilities and system learning. All participants used the same apparatus.

B.   Results

Ten questions were targeted in assessing the general usability of the UIs. The results revealed a very positive assessment regarding the usability of the UIs, but with some clear preferences (Fig.8).

Participants found the Kinect 2 system generally easy to use, not very complex and they considered that they did not need to learn many things before starting to use it, they found it consistent and not cumbersome and that they did not need any technical assistance. Additionally they felt very confident in using the UI and they were willing to use it frequently. Overall they had a pleasant experience using the Kinect 2 UI.

When participants used the Hydra system that used navigation in all 3 dimensions, there were noticeably different results. The perception of complexity was much greater, although still in the range of neutrality. There were lower marks in ease of use and frequency they would like to use it, as well as confidence in the use, quickness to learn and integration of functions. The Hydra system scored higher than the Kinect 2 in complexity,

inconsistency, need of technical support, inconsistency, cumbersomeness, and the need to learn more before using. The Hydra UI therefore showed clear indication of being less desirable for interaction.

The next part of the evaluation focused on the systems' capabilities (Fig.9).The aim was to test the systems' speed and reliability along with other technical characteristics. The results were largely neutral along the scale regarding the speed and the reliability of the UIs with reliability between systems being even.

The major difference was in the perceived speed between UIs. The Kinect 2 scored higher than the Hydra by nearly one full point. But the Kinect 2 also scored nearly half a point higher in noisiness, and slightly lower in ability to correct mistakes. It also scored lower in its appropriateness for all users.

The last part of the evaluation focused on aspects related to learning the UIs (Fig. 10). Participants felt that the Hydra UI did not need a lot of effort to be learnt and they could operate it very easily as compared with the Kinect 2. This difference was significant.

On all other aspects the two UIs were on more equal footing. Users felt they could remember the commands and they could perform the tasks in a straightforward manner. A more neutrally marked item of both UIs was the messages on the screen. The participants felt that they were neither helpful nor unhelpful.
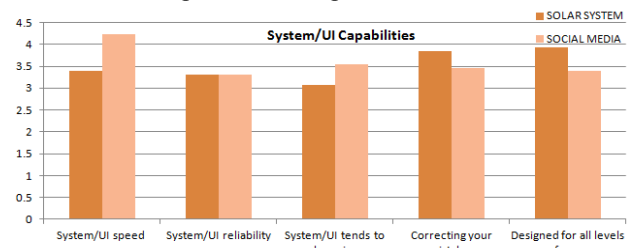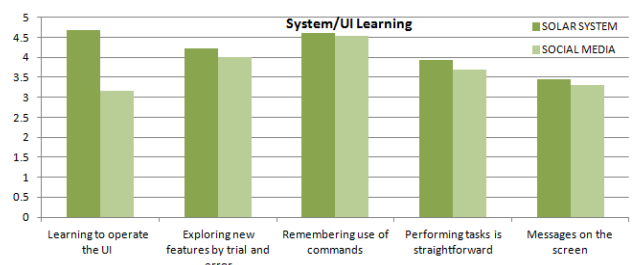

Figure 9.   System/UI Capabilities

Figure 10.   System/UI Learning

VII. CONCLUSIONS

The usability evaluation of the UIs revealed some very positive results. Participants in general found them easy to use, not complicated and they thought they were consistent and did not require a lot of effort to be learned. However, the participants had experience of such UIs and that had also affected their perceptions. Furthermore, they found the technical capabilities of the UIs very acceptable and the demands for learning the system very easy.

There were some surprising results around the Kinect 2 compared to the Razer Hydra for UI efficacy. The Kinect 2 scored better in every aspect of analysis for usability. This would merit further study to determine if a PUI with a 2 or 2.5 dimensional perceptual composition is more appropriate than one with 3 dimensions in perceptual composition.

This seems to bear out the findings of Jacob and Sibert where perception of performance on a more limited, less integrated system such as the mouse scores higher with users in perceived efficacy than that of an integrated 3 dimensional device such as the motion trackers they used which are comparable to our Razer Hydra.

The next steps of the work will be to evaluate the effectiveness of the varieties of these immersive environments with new control systems against traditional GUI interface and other PUI interfaces.

The obvious progression of this would be to enable both the Kinect 2 and Hydra systems for mouse control in order to do repeated measures designs for all resultant combinations. In this way the integrated 3d control to task mapping system on the Hydra could be compared to a Hydra system with the mouse operating 2d controls of a 3d environment. This would be similar to the Kinect interface in which the environment is 3d, but has no significant impact on the interaction (sometimes called 2.5d).

This will compare an integrated 3 dimension perceptually composed system to the same system but with a control to task mapping that is essentially a 2d screen translation of the full 3d immersive world. This would seek to answer the question of what an integrated 3rd dimension would bring to an immersive information retrieval environment by studying perception, sentiment, task to control mapping efficacy, interface efficacy via Fitts Law, Hick Hyman Law, and information retrieval efficacy.

A further experimental extension of this based on Jacob and Sibert could also be enabled where the mouse scroll wheel operates a separate (not integrated) 3rd dimension that enables a full 3d comparison between the original integrated 3d system and a separated 3d control system. This would establish a solid comparison of control to task mapping for perceptually different dimension compositions (integrated vs. separate).

The Kinect 2 comparison will merely compare between two 2d (overlaid on 3d immersive environment) control systems, one with a perceptual user interface (Kinect 2), and one with an older graphical user interface (mouse). Again, the perception, sentiment, task to control mapping efficacy, interface efficacy via Fitts Law, Hick Hyman Law, and information retrieval efficacy would be evaluated for comparison purposes. We could determine what advantages, if any, the new PUIs hold over GUIs in specific and general instances, and inform our developing PUI comparison framework.

This way the two PUIs, Kinect 2 and Hydra, can start to be assessed from the above findings for areas of comparison to start to build the PUI comparison framework. It is intended that once a framework for testing and comparing PUIs is established, the framework will be disseminated in a further paper to establish its validity in the first instance.

It is expected and hoped that there will be take-up by other researchers, and to this end there will be an initiative to evaluate our work. Our further expectation is that large scale projects such as REVERIE [8] will benefit from this, and the elements they have begun to develop could be quantified for general comparison and promotion of valuable qualities.

As example of interface innovation, our expectation is to be able to quantify the benefits of eye tracking for easier highlighting, voice command activation for selecting highlighted items, LazyNav from REVERIE [8] for pov and navigation, and discrete finger tracking for arm interaction [4, 5].

REFERENCES

[1] M. Turk, "Perceptual user interfaces," workshop available at: http://www.cs.ucsb.edu/~mturk/pubs/TurkEC-NSFWorshop.pdf , 2004, [Accessed 18 August 2015].

[2] H.J. Shin, J. Lee, S.Y. Shin, M. Gleicher, "Computer puppetry: an importance-based approach," ACM Transactions on Graphics (TOG), vol. 20 no. 2, 2001, pp. 67-94. http://dx.doi.org/10.1145/502122.502123

[3] D. Sturman, D. Zeltzer, "A design method for whole-hand human-computer interaction," ACM Transactions on Information Systems, vol. 11 no. 3, 1993, pp. 219-238. http://dx.doi.org/10.1145/159161.159159

[4] A. Mazalek, S. Chandrasekharan, M. Nitsche, T. Welsh, P. Clifton, A. Quitmeyer, F. Peer, F. Kirschner, D. Athreya, "I'm in the game: embodied puppet interface improves avatar control," In proceedings of the ACM international conference on Tangible, Embedded, and Embodied Interaction (ACM TEI '11), 2010, pp. 129-136.

[5] A. Mazalek, S. Chandrasekharan, M. Nitsche, T. Welsh, P. Clifton, A. Quitmeyer, F. Peer, F. Kirschner, "Recognizing self in puppet controlled virtual avatars," In proceedings of the 3rd international conference on Fun and Games, 2010, pp. 66-73. http://dx.doi.org/10.1145/1823818.1823825

[6] R. Chua, D.J. Weeks, D. Goodman, "Perceptual-motor interaction: some implications for human-computer interaction," The human-computer interaction handbook, L. Erlbaum Associates Inc. Hillsdale, NJ, USA, 2003, pp. 23-34.

[7] R.J.K. Jacob, L.E. Sibert, "The perceptual structure of multidimensional input device selection," In proceedings of the ACM Human Factors in Computing Systems Conference(ACM CHI '92), 1992, pp. 211-218.

[8] E. Guy, P. Punpongsanon, D. Iwai, K. Sato, T. Boubekeur, "LazyNav: 3D ground navigation with non-critical body parts," In proceedings of the 2015 IEEE Symposium on 3D User Interfaces, 2015, pp. 43-50.

[9] R.A. Bolt, E. Harranz, "Two-handed gesture in multi-modal natural dialog," In proceedings of the ACM Symposium on User Interface Software and Technology (ACM UIST'92), 1992, pp. 7-14. http://dx.doi.org/10.1145/142621.142623

[10] I. S. MacKenzie, Input devices and interaction techniques for advanced computing. In W. Barfield, & T. A. Furness III (Eds.), Virtual environments and advanced interface design, pp. 437-470. Oxford, UK: Oxford University Press 1995.

[11] Microsoft Kinect UI Guidelines: [Accessed 18 August 2015]. Available at: https://www.google.co.uk/webhp?sourceid=chrome-instant&rlz=1C1LENP_enUS500GB500&ion=1&espv=2&ie=UTF-8#q=microsoft%20kinect%20ui%20guidelines

[12] W.R. Garner, The proceing of information and structure, Lawrence Erlbaum, Potomac, Md. 1974

[13] Facebook Simulator Software: [Accessed 18 August 2015]. Available at: http://simitator.com/generator/facebook/status

[14] I. S. MacKenzie, "Fitts' law as a research and design tool in human- computer interaction", Journal on Human-Computer Interaction, vol 7, (1), pp. 91-139, 1992 http://dx.doi.org/10.1207/s15327051hci0701_3

AUTHORS

M. Mentzelopoulos is a senior lecturer with the University of Westminster 115 New Cavendish Street, London W1W 6UW UK. He is a member of the Computational Vision and Imaging Technology (CVIT) research cluster in Westminster University. His research background is in Virtual Reality, AI for games, Serious Games, Cognitive Theory in Games, Pervasive Computing, Immersive Technologies for HCI, Video and Image Analysis and Pattern Recognition. (e-mail: mentzem@westminster.ac.uk).

J. Ferguson is a researcher with the Serious Games at Westminster University, New Cavendish Street, London W1W 6UW UK. His background is in animation and motion capture for video games, virtual reality and pervasive computing (e-mail: j.ferguson@westminster.ac.uk).

A. Protopsaltis is a senior researcher with the Institut für Lern-Innovation at Friedrich-Alexander-Universität Erlangen-Nuremberg, Schlossplatz 4 91054 Erlangen Germany. His background is in Cognitive Science, Serious Games, and ICT and Education, information processing theory and comprehension (e-mail: aprotopsaltis@gmail.com).