Low-cost, room-size, and highly immersive virtual reality system for virtual and mixed reality applications

Adrián Borrego[1], Jorge Latorre[1], Roberto Lloréns[1,2], Enrique Noé[2]
[1]Instituto Interuniversitario de Investigación en Bioingeniería y Tecnología Orientada al Ser Humano. Universitat Politècnica de València. Valencia, Spain.
[2]Servicio de Neurorrehabilitación y Daño Cerebral de los Hospitales NISA. Valencia, Spain.
rllorens@labhuman.com

Emily A. Keshner[3]
[3]Department of Physical Therapy, Temple University. Philadelphia, PA, USA.

***Abstract*—Providing realistic stimuli has traditionally posed a challenge for virtual reality researchers. Most efforts have been done with visual information. Different technical solutions, such as head mounted displays and cave systems, have been presented through the years improving the state-of-the-art instrumentation. In recent years, an increasing interest in virtual reality for rehabilitation has given rise to low-cost devices with impressive performance. We created a room-size immersive setting with low-cost technology that can be used to simulate large virtual worlds and explored by walking in the physical world.**

Keywords—virtual reality; virtual rehabilitation; head mounted display, CAVE

I. INTRODUCTION

The classical definition of virtual reality (VR) involves real time simulation and interaction through one or multiple sensory channels [1]. The Milgram continuum illustrates the extent of artificial or synthetic components that are present in the environment (Fig. 1). This scale ranges from the real or physical environment, where all the stimuli are real, to a complete virtual environment (VE), where all the stimuli are artificial.

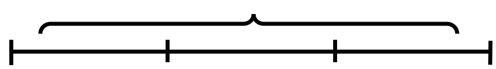

Fig. 1. Milgram continuum.

Since the visual channel provides the greatest contribution to human perception, the vast majority of VR systems have focused on visual stimulation. However, providing reliable and consistent visual stimuli in mixed and virtual reality applications has posed a technological challenge for years, due to the complex physiological properties of the human sense of sight. This sense has a field of view (FOV) that covers approximately 180º horizontally and 120º vertically in binocular vision, with an overlap of 120º in the horizontal axis [2], has an illumination range of about ten orders of magnitude [3], and can resolve a separation of about 0.5 minutes of arc [4].

Different technological solutions have been proposed in the past to provide visual stimulation for mixed reality and VR applications. The first immersive device was presented in the late sixties [5]. It consisted of a binocular head mounted display (HMD) that provided visual information to both eyes through small displays that were located in front of each eye. Years later, in the early 1990s, the first LCD-based HMD, the VPL EyePhone (VPL Research, Palo Alto, CA) was commercially available. The device had a FOV of 106x75º, a resolution of 720x480 pixels, a weight of 2.4 kg, and cost $49,000. Also in this decade, another solution was presented. The cave automatic virtual environment (CAVE) covered the whole human FOV introducing the users in a room-sized cube and projecting the VE on the walls of the cube [6]. The first CAVE system threw stereo full-color workstation fields (1280x512 pixels) at 120Hz onto the screens, and cost more than $100,000.

Through the years, technological advances have favored the appearance into the market of new versions of HMDs and CAVEs, providing better performance at lower prices. The increasing interest in VR has been recently promoted by the low-cost HMD Oculus Rift (Oculus VR®, Menlo Park, CA). The Oculus Rift has a 7" LCD screen that covers 117.4x112.5°, a resolution of 1280x800 (640x800 per eye), a weight of 380g, and costs $300. The latest version of this HMD, the Oculus DK2, has recently improved the resolution of the display to 1920x1080 pixels (960x1080 pixels per eye), together with other parameters.

Most VR settings that involve HMDs or CAVEs either require that users maintain a stationary position while interacting with the environment, or only allow them to perform limited movements such as moving one step sideways or forward/backward. The displacement in the VE is carried out by different representations, generally by using joysticks or flysticks, which can affect immersion and presence [7]. For some rehabilitation applications, it may be beneficial for users to explore a larger VE by walking. Even though different solutions can be used for motion tracking in a large space [1], they would likely result in expensive solutions not affordable by end users.

The objective of this paper is to present a low-cost solution that provides immersive audiovisual stimulation for mixed and virtual reality applications in room-sized environments.

II. Methods

A. Hardware

The VR setting presented in this paper consists of an HMD, a Bluetooth headset, three RGB cameras, a pattern of markers fixed to the ceiling, and a computer (Fig. 4). The components of the computer consist of an 8-core Intel® Core™ i7 Haswell, 8 GB of RAM, and a NVIDIA® Geforce® GTX 860M with 2GB of GDDR5, and cost $849. A description of the other components is provided below.

Visual stimulation and orientation tracking: This two requirements are satisfied by a single device. An HMD shows the visual content to the users and provides information about the orientation of the head in the yaw, pitch, and roll coordinate frames. In the experimental setting presented here, both the Oculus Rift and the Oculus DK2 can be used (Fig. 2).

a) b)

Fig. 2. HMD used in the system: a) Oculus Rift; b) Oculus DK2.

The characteristics of both devices are described in TABLE I.

TABLE I. Comparison between versions of the Oculus

Parameter	Oculus Rift	Oculus DK2
Screen size	7" LCD screen	5.7" OLED (PenTile)
FOV	117.4x112.5°	100°
Refresh rate	60 Hz	75 Hz
Persistence		2 ms
Resolution	1280x800 (640x800 per eye)	1920x1080 (960x1080 per eye)
Orientation tracking	Gyroscope, accelerometer, magnetometer @ 1KHz	Gyroscope, accelerometer, magnetometer @ 1KHz
Positional tracking	No	Near infrared CMOS Sensor @60 Hz
Weight	380 g	440 g
Cost	$300	$350

Audio stimulation: A Bluetooth headset is used to reproduce the audio information. A wireless device was selected to avoid motion limitations of wires. The Kinivo BTH220 (Kinivo™, Bellevue, WA), which costs $20.99 and weighs 71 g, is used in our experimental setting (Fig. 3).

Fig. 3. Kinivo BTH220

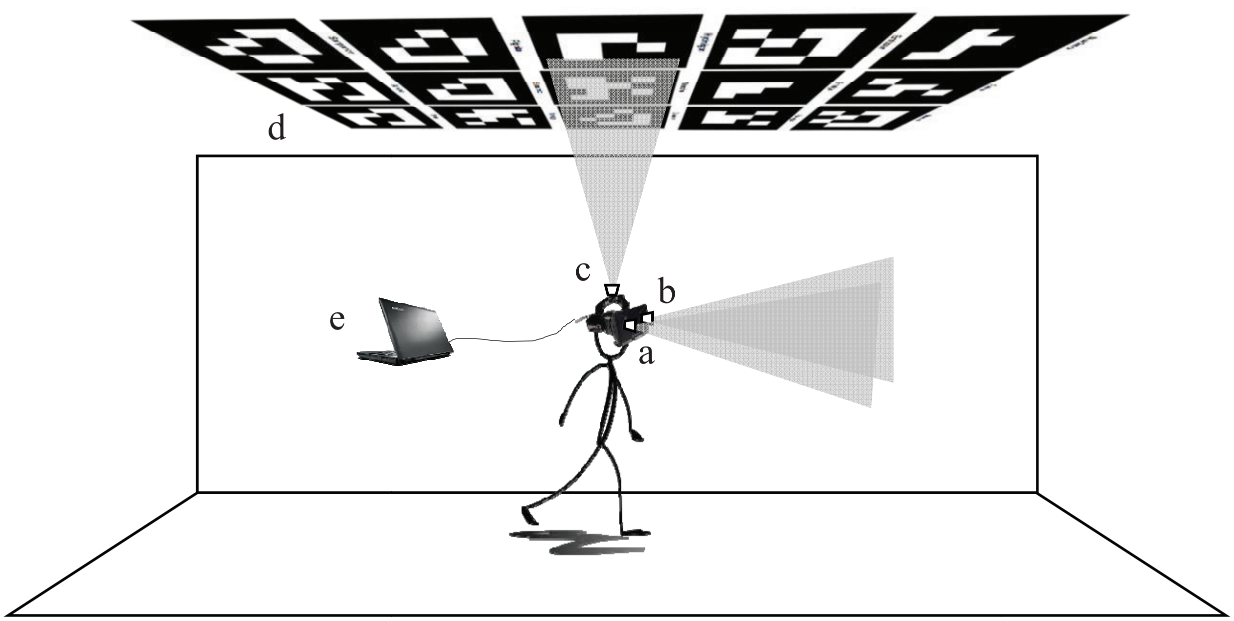

Fig. 4. Diagram of the proposed virtual reality system: a) Oculus Rift; b) Ovrvision; c) Playstation Eye camera; d) pattern of fiducial markers; e) computer.

Physical world capture: The goal of this component is to capture the physical world for mixed reality applications. The Ovrvision device (Shinobiya.com Co., Ltd Wizapply, Osaka, Japan) is used (Fig. 5). The Ovrvision consists of two RGB cameras separated 5 cm, simulating the interpupillary distance. This device has a resolution of 1280x480 pixels (640x480 pixels each camera), a frequency rate of 60 Hz, low-latency data transfer, a weight of 55 g, and costs $159. The lenses of the camera have a focal length of 2.2 mm and a FOV of 135º in diagonal that allows focused distances larger than 8 cm and provides stereovision at distances larger than 30 cm.

Fig. 5. Attachment of the Ovrvision to the Oculus DK2.

Positional tracking: This component provides feedback about the position of the user with three degrees of freedom by determining the x, y, and z coordinates of the user. A marker tracking solution was adopted to maintain the low-cost goal. The solution consists of a pattern of fiducial markers fixed to the ceiling, and an RGB camera attached to the upper side of the HMD that points at the markers (Fig. 4). A PlayStation®Eye Camera (Sony® Corporation, Tokyo, Japan) is used (Fig. 6). The camera is configured to capture standard video with a frame rate of 75 Hz at a 640×480 pixel resolution. The body of the camera was dismantled to reduce weight, and only the camera is used. The final weight of this device is 75 g. An additional lens (4.3 mm, 70º FOV) was mounted on the camera to improve the image quality. The camera and the lenses cost $10 and $15, respectively. The pattern of markers consists of 17x26 fiducial markers, 18x18 cm each, separated 4 cm, thus covering an area of 3.78x5.76 m. The pattern was printed on a vinyl sheet and was fixed to the ceiling. The impression on the vinyl cost $350.

Fig. 6. PlayStation®Eye Camera.

It is important to highlight that even though the Oculus DK2 includes an external infrared camera to provide positional tracking, it is only capable of tracking short displacements of the head and users must be in front of the camera. Since the requirement of the system was to cover a room-size space, this tracking capability was discarded.

B. *Software*

The VE was generated using Unity 3D (Unity Technologies ApS, San Francisco, CA). A project was programed including the plugins of the Oculus devices, the Ovrvision, and the PS3 camera, and the positional tracking library. The ArUco library, a minimal library for augmented reality applications based on OpenCv, is used [8]. Even though this library provides information of both position (x, y, and z coordinates) and orientation (yaw, pitch, and roll), only positional tracking was considered. As previously stated, the orientation of the head, and thus the orientation of the camera, is provided by the HMD.

III. Conclusions

This paper presents a low-cost system to provide immersive audiovisual stimuli in a room-size environment for mixed and virtual reality applications.

The setting described here covers an area of almost 24 m^2, which should be enough for those applications that need to simulate room-size spaces (Fig. 7). However, the area covered by the system can be extended adding more markers to the pattern (with a maximum of 1024 markers) or changing the size of the markers, and so the lens of the positional tracking camera. It is important to highlight that the system allows exploration of the VE without the need of navigation devices or representations. Users can move in the VE as they would in the physical world, which we expect would increase immersion in the virtual world. Preliminary studies of the accuracy of the positional tracking, measuring the accuracy of the estimated position of the HMD at a height of 175 cm, showed jitter [1] values of 3 mm, which can be considered accurate enough for most applications.

One important feature of the described system is the inclusion of physical world information in the VE, which makes it suitable for mixed reality applications. Seeing our own body can increase body awareness and modulate our actions in the VE [9]. This fact could be especially relevant for those applications interested in measuring motor responses to virtual stimuli, such as virtual motor rehabilitation applications.

Even though the design of the experimental setting took into account minimizing the weight of the head mounted device, the final weight was 641 g (TABLE II.). Even though it is lighter than earlier systems, this weight can still be considered high for some applications.

TABLE II. WEIGHT DISTRIBUTION OF THE HEAD MOUNTED DEVICES

Component	Weight
Visual stimulation and orientation tracking Oculus DK2	440 g
Audio stimulation Kinivo BTH220	75 g
Physical world capture Ovrvision	55 g
Positional tracking Playstation Eye Camera	71 g
Total weight	641 g

With regards to the final cost, the setting presented here cost $1753.99 (TABLE III.). The computer is the largest portion of the cost at almost half of the total expenses (48.4 %). The cost of the Ovrvision should be removed for complete VR applications, where information about the physical world is not needed. This cost should make the system affordable to many research groups.

The experimental setting presents some limitations that should be highlighted. First, the distance between the cameras of the Ovrvision is fixed at 5 cm, which is less than the reported interpupillary distance of 6.47 cm for men and 6.23 cm for women [10].

TABLE III. COST DISTRIBUTION OF THE EXPERIMENTAL SETTING

Component	Cost
Visual stimulation and orientation tracking Oculus DK2	$350
Audio stimulation Kinivo BTH220	 $20.99
Physical world capture Ovrvision	 $159
Positional tracking Playstation Eye Camera Lens Vynil	 $10 $15 $350
Computer	$849
Total cost	$1753.99

Fig. 7. User interacting with the proposed virtual reality system. Top: Experimental setting. a) Oculus DK2; b) Ovrvision; c) Playstation Eye camera; d) Kinivo BTH220; e) pattern of fiducial markers; f) virtual environment that is being displayed. Bottom: Snapshot of the virtual environment (grocery store) where user is immersed.

In addition, the optics of the Ovrvision are different from the eyes. These factors alter the visualization of the physical world and could have a strong impact on the users' performance. Second, besides the weight of the head mounted devices, it is important to highlight that all components are connected to the computer through wires. Consequently, either an experimenter is needed to carry the computer and to handle the wires to avoid tangles, or the user has to carry the devices, for instance, in a backpack. Both options are expected to have an impact on the comfort of the solution. Finally, the lighting condition of the physical environment should be controlled so that the markers are correctly captured by the positional tracking camera. It should be noted that the pattern of markers are likely to occlude the ceiling lighting. In that case, the markers should be illuminated from below by homogeneous and indirect white light. Future studies should assess the perception of the body in mixed reality environments, explore how this perception might modify movement patterns, evaluate the comfort of the proposed setting, and determine how these factors affect immersion and presence. However, preliminary testing has indicated promising performance of the experimental design.

ACKNOWLEDGMENT

This study was funded by the Ministerio de Economía y Competitividad of Spain (Project NeuroVR, TIN2013-44741-R).

REFERENCES

[1] Burdea, G. and P. Coiffet, *Virtual Reality Technology*. 2nd ed. 2003, New Jersey: John Wiley & Sons, Inc. 464.

[2] Kalawsky, R., *The Science of Virtual Reality and Virtual Environments*. 1993: Addison-Wesley Longman Publishing Co., Inc.

[3] Jain, A.K., *Fundamentals of digital image processing*. 1989: Prentice-Hall, Inc. 569.

[4] Bruce, V., *Visual perception : physiology, psychology and ecology / Vicki Bruce and Patrick Green*, ed. P.R. Green. 1990, Hove: Erlbaum.

[5] Sutherland, I.E., *A head-mounted three dimensional display*, in *Fall Joint Computer Conference*1968: San Francisco. p. 757-764.

[6] Cruz-Neira, C., D.J. Sandin, and T.A. DeFanti, *Surround-screen projection-based virtual reality: the design and implementation of the CAVE*, in *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*1993, ACM: Anaheim, CA. p. 135-142.

[7] Patrice L. Weiss, R.K., Uri Feintuch, et al., *Virtual reality in neurorehabilitation Textbook of Neural Repair and Rehabilitation*. 2006: Cambridge University Press.

[8] Garrido-Jurado, S., et al., *Automatic generation and detection of highly reliable fiducial markers under occlusion.* Pattern Recognition, 2014. **47**(6): p. 2280-2292.

[9] Kilteni, K., I. Bergstrom, and M. Slater, *Drumming in immersive virtual reality: the body shapes the way we play.* IEEE Trans Vis Comput Graph, 2013. **19**(4): p. 597-605.

[10] Gordon, C.C., et al., *1987–1988 Anthropometric Survey of U.S. Army Personnel: Methods and Summary Statistics*, 1989, U.S. Army Natick Research, Development and Engineering Center: Natick, MA.