https://doi.org/10.2352/ISSN.2470-1173.2017.3.ERVR-094

CAVE versus Head-Mounted Displays: Ongoing thoughts

Daniel R. Mestre, Aix Marseille Univ, CNRS, ISM/CRVM, Marseille, France

Abstract

In a short period of time, virtual reality has taken over the media, tending to promote that idea that it is a new technology. In fact, all started as early as the seventies and eighties, for portable devices (e.g. Head-Mounted Displays) as well as for complex and large devices (CAVEs). In this paper, we try to put these different systems in perspective, and to show the interest of comparing them in an experimental approach.

Introduction

If you look at the news and/or social network sites, you might think that the age of the HMD (for Head-Mounted Display) has come, and that it is going to be the ultimate display Ivan Sutherland was dreaming of in 1965 [1]. Well, for one thing, HMD are no new story. Back to 1965, lucky ones could test the "Sword of Damocles (figure 1). Since then, it is certainly true that technological power has taken over the wireframe cube rotating in the viewer's field of view. However, for this device and the idea of immersing the viewer in a surrounding virtual environment, Sutherland is often coined as the "inventor" of Virtual Reality. For him, the purpose of virtual reality was to look at the (visual) display as a "window" to a realistic virtual environment.

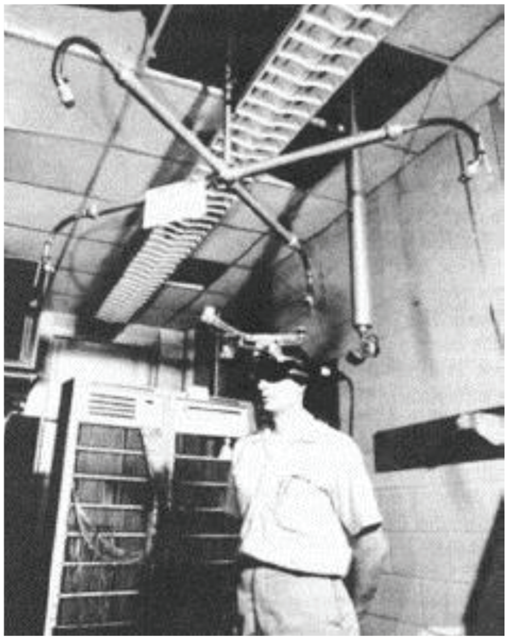

Figure 1. The Sword of Damocles [1]. Two cathodic tubes cover the subject's eyes; his motions are acquired by a robotic arm, to insure real-time adjustments of 3D images. This is considered as the first Virtual Reality system.

Since then, many advances have been made, including developments by NASA Ames Research Center in the Eighties [2, 3], or by people around Fred Brooks [4], at Chapel Hill (such as Henry Fuchs, and the recent initiative for telepresence, through the BeingThere International Center). The "Virtual Environment Workstation" included a LCD display and a Polhemus head tracker and represented, on that respect, a significant technological progress, as compared to the "Ultimate Display" (figure 2). We are today getting closer to that "embodied" vision, which is not far from the vision of a totally immersive cinema, stimulating all the senses, by Morton Heilig [5] or the proposition of Virtual Reality as a "mind-expanding" feature, by Myron Krueger [6].

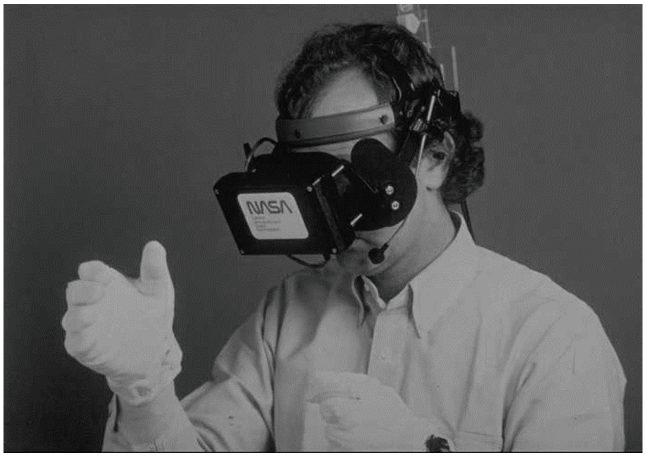

Figure 2. The NASA's First HMD. Note the first "dataglove", used for interacting with the virtual environment.

It would be pointless to enumerate all the HMDs that were invented and/or commercialized. For an historical view, the reader might refer to [7].

To make a long story short, we are now facing modern HMDs, like the Oculus Rift or the HTC Vive, that come to the (potentially) mass-media market, with relatively low prices and good spatio-temporal resolution. However, see below.

Of the other side (of the VR arena) is the CAVE system [8]. The main purpose of this multi-screen device is to stimulate the entire visual field of the subject (see figure 3), while insuring real-time interactivity, using motion-tracking cameras.

The CAVE has evolved as well, and a few attempts were made to develop "low-cost" CAVEs [10]. An interesting recent development is the StarCAVE [11], developed in collaboration between EVL in Chicago and UCSD. This system optimizes visual and auditory immersion, as compared to the original CAVE concept.

Figure 3. A subject inside a CAVE system, consisting in 4 screens (3 vertical "walls" and a "floor". This is the most common type of CAVE, using DLP stereoscopic projectors. A few six-sided CAVES exist, favoring a more complete immersion. See for instance [9].

From this quick review, it is clear that both HMD and CAVE systems have coexisted from the "birth" of Virtual Reality (that is 1965 and not merely 2016), and continue to evolve over the years, along with technological advances, including computing and rendering power as well as material cost.

Today, some will claim that the huge difference in price between a HMD (a thousand dollars) and a Cave (a couple hundred thousand dollars) is a potential "killer" factor. However, this simple argument is (perhaps) valid for the mass market of video gaming. It however fails to consider all the factors involved in the respective utility of different devices. In the following lines, we will present some arguments, open for future discussion.

Pros and Cons

First of all, the main serious argument (and valid as of today's technology) is the fact that the CAVE is also a collaborative environment, whereas a HMD isolates the user from his/her immediate surroundings, including other humans. Initiatives are now been made to go around this limitation of HMDs, including Kinect or Leap Motion devices, used to reproduce a sketchy picture of the real environment.

However, it remains that a CAVE is suited for collaborative work, and acts in this sense as an immersive Computer-Aided Design device. In this sense, the Fraunhofer IAO Institute has developed a complete argumentation on that point [12], around the concept of Immersive Engineering. One limitation of this argumentation is that, usually, only one person is tracked in a CAVE, thus having the correct point of view on the environment. This is true, but recent input from industrial partners suggest that CEOs strongly prefer the CAVE solution, even with this one-person imitation, as compared to the HMD (including hygienic concerns).

Talking about immersion, an HMD seems to be better, since the virtual environment is visible all around the subject, which is not the case in a "classical" four-sided cave, in which the ceiling and back-screen are "missing". This means that the subject loses immersion whenever she looks up (expect in a five-sided cave) or back. As mentioned, there are a few six-sided cave. Besides their cost, they are technically complex and might cause claustrophobia, not mentioning ventilation problems.

On this aspect, the HMD seems to be a winner. However, this notion of "everywhere" immersion is misleading, since the subject had to turn his head to see the virtual environment. This is related to two problems that have different aspects, depending on the device.

First, the extent of the instantaneous field of view is commonly around 100 degrees (less in previous generations) in a HMD. This is significantly inferior to the human actual field of view, being around 190 degrees in the horizontal dimension and 120 in the vertical dimension. On this aspect, the CAVE field of view is close to these values (with the limitations mentioned above). There has been a few HMDs with larger fields of view. The Sensics piSight system, using a configuration of 12 screen panels per eye had a field of view up to 150 degrees. It is however quite difficult to calibrate (geometrically and in terms of colorimetry and luminance). It also requires a cluster of computers to run, which disqualifies it as a "low-cost" HMD. It is however worthy to mention that efforts are still being to investigate the usability of high-end HMDs, in an industrial and/or scholar context.

Secondly, with the HMD, having to turn one's head to "see the world" means graphics computing, which means temporal delay and latency. It is well known now that latency means 1) a perceived non stability of the virtual environment, which leads to breaks in presence [13] and 2) potential cybersickness, due to sensorimotor incongruences and/or a visuo-vestibular sensorial conflict [14]. Brooks [4] mentions that professional jet pilots are able to mention latencies as low as 50 msec. This is significantly inferior to classical "real-time loop" in a CAVE system, running at 60 Hz, in which the overall latency is never inferior to 3 times the base temporal base of 17 msec, hence a lower threshold of around 50 msec. In fact, the latency is always between 50 and 100 msec in a CAVE. This is correct to enable acceptable visuo-motor control (a number of previous HCI studies has shown that latencies above 100 msec significantly impair performance, while latencies above 300 msec make visuomotor coordination almost impossible.

On this problem, recent developments by Oculus or Vive have made significant progress. In particular, Oculus announces an overall latency of 26 msec. On this point, these HMDs have a clear advantage over a CAVE. They use their own motion tracking system and optimize graphics rendering. It will be interesting to see whether progress can be made in a CAVE, notably by getting rid of PC clusters with modern graphics boards. It remains that one need to test if the graphics "tricks" to temporally optimize rendering in HMDs is or not susceptible of distorting precise perception of virtual environments. Here again, the target market is of importance, since a gamer does not pay much attention to graphics shortcuts and approximation, while this aspect is decisive in a "serious" application searching for valid conclusions from virtual reality experiments.

Finally, a few words about spatial and temporal resolution. First, concerning spatial resolution, Michael Abrash, at Oculus Connect 2014, mentioned that we are far from matching the resolution power of the human eye. Today, our favorite HMDs offer the same number of pixels per eye: 1200 x 1080. High-end CAVE systems deliver over 16 million pixels par screen, for a total of 100M pixels [9]. However, the number of pixels does not really make sense, since the angular size of a pixel, which refers directly to the resolution power of the eye, depends on the distance of the eye to the projection screen. Moreover, when Michael Abrash mentioned the huge pixel resolution necessary to match the human eye, he assumed that we need optimal resolution at any moment all over the display screen. This is the default option, because we do not know when and where the eye will fall. However, we also know that the resolution of our visual system is not the same across the visual field, being highest at the fovea and decreasing dramatically in peripheral vision (for instance, refer to the cortical magnification factor [15]). Confronted with this fact, vision scientists have long been trying to develop gaze contingent displays [16]. This manipulation of the rendered image requires real-time acquisition of the position of gaze in the virtual environment, which, by itself, poses a number of technological problems. However, advances have been made, including gaze capture inside a HMD. The real problem with gaze contingency is in the temporal domain. Saccades are (more or less) unpredictable and incredibly fast (over 600 degrees/sec peak velocity). In the current state of the technology, this would require tracking latencies well below what is achievable.

Now, concerning temporal resolution: Most CAVE systems run at 60 frames per second, while "modern" HMDs run at 90 Hz. Referring to early standards of VR, saying that 20 Hz was enough, we are now in a more comfortable zone of smooth motion perception. However, it is illusory to say that 90 Hz is always enough to enable the perception of continuous motion. In fact, systematic studies revealed that there an upper displacement limit of a given pixel from one frame to the next, for a given frame rate, above which the continuity of motion is lost [17]. In our own experiments [18], using SGI hardware at 75 Hz, we found that it was not possible to investigate retinal speed above around 30 degrees/second. In other words, as compared to the analogical nature of our visual system, the digital nature of computer graphics poses serious challenges and more power/resolution remains to be developed.

Finally, we will evoke quickly the accommodation-vergence issue in virtual environments. The fact that we use flat displays (being computer screens, CAVE walls or HMDs' OLED screens) results in the fact that accommodation is fixed on the screen. We will not discuss here the fact that this might cause serious ophthalmological problems, such as myopia for future generations using tablets from birth. More specifically, in the case of stereoscopic displays, this creates a decoupling between vergence and accommodation mechanisms of human vision. Recent studies have shown that such decoupling of a natural behavior impairs visual performance and causes visual fatigue [19, 20]. This problem exists in the CAVE as well as in the HMD.

Up to now, we saw that many problems, of technological as well sensorimotor nature, exist whatever the system used, and that progress has to be made at a number of levels: oculomotor coupling, temporal and spatial resolution, system latency, extent of the field of view. We would like to turn now to short (narrative) reports of recent experiments using one system or both, to give some feedback at a more experiential level

Experimental observations

Interacting with objects

First of all, we encountered a classical problem, when studying simple reaching and tracking arm and finger movements inside a CAVE system [21]. We were trying to study the effect of vibrotactile feedback for collision avoidance, since haptics feedback remains a hard problem in virtual reality.

In fact, in a CAVE, having a force-feedback poses a perceptual problem, since the apparatus itself has (often) nothing to do inside the virtual environment. On the other hand, not having haptics feedback, does not prevent subjects to "pass through" virtual objects. Once that happens, the objects instantaneously become transparent and their "presence" is lost. Such a problem does not exist using an HMD. However, as we will see, other problems arise.

Acted affordances

In this line of research, we evaluated embodied presence in virtual environments, which is the way presence in VR can be directly observed at a behavioral level. To do so, we designed a simple experiment, in which subjects were asked to pass through an aperture of variable width. In normal (real) circumstances subjects systematically rotate their shoulders, in order to avoid "colliding" the environment [22]. We essentially observed the same behavior in a virtual environment [23], thus qualifying our approach to presence.

Figure 4. Representation of the virtual environment, with the variable width aperture and the subject rotating his/her shoulder to pass through.

We decided to replicate this experiment using a HMD (Oculus DK2), in a similar virtual environment (figure 4). In this case, one observation can be immediately made: when you put the helmet on, you no longer see your own body. We then decided to use the vision of an avatar of the subject's own body as an experimental factor. To do that, the subject was equipped by whole body markers (ART system) and a co-localized avatar of him/herself could be inserted in the visual display.

Results are straightforward. Without the self-avatar, subjects collide with the aperture borders more than half of the time (for a small aperture of 40 cm). Collisions are significantly reduced with the self-avatar [22]. Such results clearly demonstrate that the presence of a visual representation of your own body, co-localized in real-time with your own movements is decisive in enabling one to calibrate body-environment relationships.

At first view, this is a clear advantage for the CAVE, in which the vision of your own body is always present.

Clinical feedback

Another supporting argument for the crucial role of the vision of your own body comes from recent evaluations we conducted during a clinical protocol involving Virtual Reality Exposure

Therapy for patients suffering from acrophobia (fear of heights). Figure 3 (above) shows an example of one subject crossing a high bridge across a deep canyon.

In a few words, this type of exposure therapy, in which the intensity of the virtual height can be adjusted to a given patient at a given movement works remarkably well, as was previously documented [25]. However, during the therapeutic process, it was evident, from observing the patients, that the natural vision of their own body was important in the process, maybe more so in the case of acrophobia therapy, because the patients were almost always looking down.

Conclusion

Form what was sketched above, it is quite difficult to say which device is best, between a CAVE and a HMD. For both devices, there are important factors (field of view, latency, visual fatigue, cybersickness...) that remain to be fixed. These current limitations exist, at different levels, for both devices. The money argument is certainly a significant one. However, it certainly has more weight in the general public than for professional/scientific applications, in which the validity and generalization of the results obtained during virtual reality evaluations is fundamental.

One argument we finally want to put forward is the fact that a HMD isolates, at least visually, the subject from his/her surroundings. One might argue that, in the future, cluster of synchronized HMDs will be available, in which a number of subjects will be able to see and interact with each other.

However, we are not there yet. From our own experiments with HMDs, we concluded that a vision of your own body has to be brought into the HMD. Clearly, systematic comparisons between CAVE and HMD are missing. In the domain of virtual therapy, this issue is important, since patricians would rather use a portable (and low-cost) HMD rather than having to bring their patients in a CAVE facility.

We are aware of two papers trying to compare HMD and CAVE for the treatment of phobias. [26, 27]. They failed to find any difference between both systems. However, they used questionnaires of presence to try to assess the difference. As presented above, we believe that behavioral evaluations are required to obtain a more objective vision of the effects of HMD and CAVE systems.

Acknowledgments

The author wishes to thank his collaborators in the mentioned VR studies: Jean-Claude Lepecq, Cephise Louison, Eric Malbos, and the Mediterranean Virtual Reality Center team (www.crvm.eu). This on-going discussion on VR devices was initiated during a workshop in Bordeaux, in 2015, during the annual meeting of the French Association for Virtual Reality (www.af-rv.fr), co-organized with F. Guillaume (Airbus Company) and A. Solignac (Iceberg Research Group).

References

[1] I.E. Sutherland, "The 'Ultimate display", in Proceedings of the Information Processing Techniques Congress, pp.506—508, 1965.

[2] S.S. Fisher, J. Humphries, M. McGreevy, W. Robinett, “The Virtual Environment Display System,” 1986 ACM Workshop on Interactive 3D Graphics, Chapel Hill, 1986

[3] J.D. Foley, 'Interfaces for Advanced Computing" (cover story), Scientific American magazine, October 1987.

[4] https://medium.com/beyond-the-headset/in-the-beginning-how-vr-was-born-an-interview-with-fred-brooks-8c09354b0b58#.xdd3sd1rm

[5] M. Heilig, "El cine del futuro: the cinema of the future", *Presence: Teleoperators & Virtual Environments*. 1, 279-294, 1992.

[6] M. Krueger. Artificial Reality 2, Addison-Wesley Professional, 1991.

[7] http://stereo3d.com/hmd.htm

[8] C. Cruz-Neira, D. Sandin, & T. DeFanti, "Surround-screen projection-based virtual reality: The design and implementation of the CAVE", Proceedings of SIGGRAPH’ 93, 135-142, 1993

[9] http://www.vrac.iastate.edu/facilities/c6/

[10] A. Juarez, B. Schonenberg, & C. Bartneck, "Implementing a Low-Cost CAVE System Using the CryEngine2". Entertainment Computing, 1(3-4), 157-164. 2010.

[11] T. De Fanti et al. "The StarCAVE, a third-generation CAVE and virtual reality OptIPortal", Future Generation Computer Systems, 25, 2, 169-178, 2009.

[12] https://blog.iao.fraunhofer.de/dont-be-blinded-by-the-oculus-rift/

[13] B.Liebold, M. Brill, D. Pietschmann, F. Schwab & P.Ohler," Continuous Measurement of Breaks in Presence: Psychophysiology and Orienting Responses", Media Psychology, pp 1-26, 2016.

[14] S.Sharples, S. Cobb, A. Moody, J. R. Wilson, " Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display system", Displays, 29, 2, 58-69, 2008.

[15] A. Cowey & E.T. Rolls, "Huamn cortical magnification factor and its relation to visual acuity", Exp. Brain Res, 21, 447-454, 1974.

[16] A. Duchowski, "Gaze-contingent displays: A review", Cyber-psychology and Behavior, 7, 6, 621-634, 2004.

[17] K. Nakayama K, G. Silverman, "Temporal and spatial characteristics of the upper displacement limit for motion in random dots", Vision Research, 24, 4? 293-299, 1984.

[18] D. R. Mestre, G. S. Masson & L. S. Stone, "Spatial scale of motion segmentation from speed cues. Vision Research, 41, 2697-2713, 1999.

[19] S.K. Rushton & P.M. Riddell, "Developing visual systems and exposure to virtual reality and stereo displays: some concerns and speculations about the demands on accommodation and vergence. *Applied ergonomics*, 30, 1, pp.69-78, 1999

[20] D.M. Hoffman, A.R. Girshick & M.S. Banks, "Vergence – accommodation conflicts hinder visual performance and cause visual fatigue", Journal of Vision, 8, 33, pp.1–30, 2008.

[21] C. Louison, F. Ferlay, D. Keller & D.R. Mestre, "Vibrotactile feedback for collision awareness", Proceedings of the 2015 British HCI Conferencce, Lincoln, UK, pp. 277-278, 2015

[22] Warren, W.H. and Whang, S., "Visual guidance of walking through apertures: Body-scaled information for affordances," Journal of Experimental Psychology: Human Perception and Performance, 13(3), 371–383 (1987).

[23] J.C. Lepecq, L Bringoux, J.M. Pergandi, T. Coyle & D.M. Mestre, "Afforded actions as a behavioral assessment of physical presence in virtual environments. Virtual Reality, 13, 141-151, 2009.

[24] D.R. Mestre, C. Louison & F. Ferlay, "The Contribution of a Virtual Self and Vibrotactile Feedback to Walking Through Virtual Apertures", Proceedings of the International Conference on Human-Computer Interaction, Toronto, Canada, pp. 222.232, 2016.

[25] B.O. Rothbaum, L.F. Hodges, R. Kooper, D. Opdyke, J. Williford, & M. North, "Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia. Am. J. Psychiatry 152, 626–628, 1995.

[26] M. Krijn et al, "Treatment of acrophobia in virtual reality: the role of immersion and presence", Behaviour Research and Therapy, 42, 229-239, 2004.

[27] K. Meyerbröker, N. Morina, G. Kerkhof & P. M.G. Emmelkamp, " Virtual Reality Exposure Treatment of Agoraphobia: a Comparison of Computer Automatic Virtual Environment and Head-Mounted display", Annual Review of Cybertherapy and Telemedicine 2011, 51-56.

Author's Biography

Daniel R. Mestre is a senior researcher in the Institute of Movement Sciences (www.ism.univ-amu.fr) in Marseilles, France. He is Head of the Mediterranean Virtual Reality Center (www.crvm.eu) and a founding member of the Association for Virtual Reality (www.af-rv.fr). His major research interests concern active visual motion perception, psychophysical and neuropsychological approaches to sensory information processing. Virtual reality is used as an experimental tool to understand our presence to the world, in both fundamental and applied domains.