Editor: Miguel Encarnação

Bringing VR and Spatial 3D Interaction to the Masses through Video Games

Joseph J. LaViola Jr.
University of Central Florida

Virtual reality (VR) and 3D user interface (3D UI) research has been going on for more than 20 years, with limited success in being a part of people's everyday lives. Don't get me wrong; VR and 3D UIs have been useful in areas such as simulation and training, scientific visualization, medicine, and psychological rehabilitation.[1] But to say that VR and 3D UI research has significantly affected mainstream society is undeniably false. Compared to other technologies such as the mouse and WIMP (windows, icons, menus, point and click) interfaces, VR and 3D UIs have had little impact.

However, we're seeing a change in the utility of VR and 3D UIs, with concepts from these fields emerging on a large, mainstream scale. The video game industry, as it has done for real-time 3D computer graphics for the last 15 years, is bringing spatial 3D interaction into people's living rooms through rather simple, yet effective technologies such as computer vision and motion sensing. What's even more interesting is that the video game industry has tried many times before to bring VR technology to its customers and failed miserably. Devices such as the Nintendo U-Force and Sega 3D glasses are just two examples of such failures.

The question then is why have recent innovations such as the Sony EyeToy and Nintendo Wii been so successful? Here, I look at a possible answer and discuss the research opportunities presented by the latest commercial push for spatial 3D interaction in games.

3D UIs and video games

What do I mean by 3D UIs in video games, anyway? In general, 3D UIs involve input devices and interaction techniques for effectively controlling highly dynamic 3D computer-generated content,[2] and there's no exception when it comes to video games.

A 3D video game world can use one of three basic approaches to interaction. The first maps 2D input and button devices, such as the keyboard and mouse, joysticks, and game controllers, to game elements in the 3D world. This is basically the traditional approach; it's how people have been interacting with 3D (and 2D) video games since their inception many years ago.

The second approach simulates the real world using replicas of existing devices or physical props. Common examples include steering wheels, light guns, and musical instruments (for example, the guitar in Guitar Hero). These devices don't necessarily provide 3D interaction in the game but do provide 3D input devices that enable more realistic gameplay.

The third approach is true spatial 3D tracking of the user's motion and gestures, where users interact in and control elements of the 3D gaming world with their bodies. This control can come through vision-based devices such as the Sony EyeToy and motion-sensing devices such as Wiimotes.

For bringing 3D UIs and VR to the masses, the second and third approaches hold the most promise and are primed to take video games to the next level of innovation.

Historical perspective

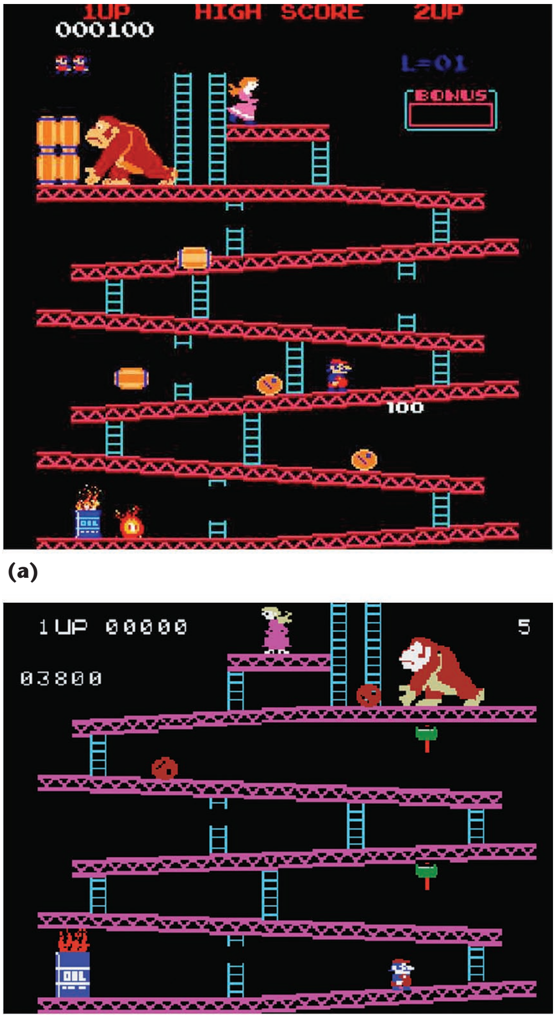
Briefly examining past video game trends will provide insight into why 3D UIs are here to stay and what the future holds for them.

Video arcades and game consoles

I've been playing video games in arcades, on consoles, and on PCs for almost 30 years. So, I've seen how technology has progressed from the days of Pong and Space Invaders to Grand Theft Auto IV and Halo 3.

The video game industry began in the mid-'70s. During these early days, if you wanted to play video games, you had two choices: go to an arcade or buy a game console such as the Magnavox Odyssey, which debuted in 1972. (PCs became popular in the 1980s but were expensive compared to game consoles.) Arcades were much more popular than game consoles because you could pack more graphics, sound, and game play into a large, deco-

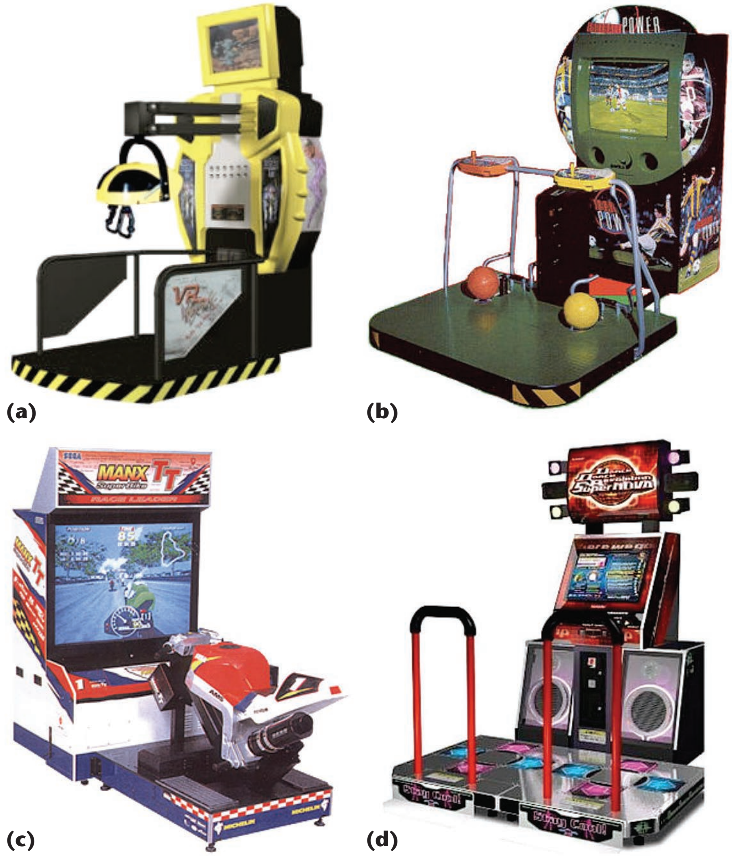
 Published by the IEEE Computer Society



(a)

(b)

Figure 1. Donkey Kong: (a) the arcade and (b) ColecoVision game console versions. Game console graphics quickly approached the quality of arcade games.



(a) (b)

(c) (d)

Figure 2. Arcade games that used more realistic input strategies: (a) BeachHead, (b) Football Power, (c) Manx TT, and (d) Dance Dance Revolution. Such innovative interfaces helped arcade games compete with PCs and game consoles.

rated, wooden cabinet than you could into a relatively small console.

This trend continued through the ’80s as arcade game graphics and sound became more sophisticated. However, game consoles weren’t far behind, as systems such as ColecoVision, the Sega Master System, and the Super Nintendo Entertainment System emerged. When I first played Donkey Kong on ColecoVision, I was amazed at how close it was to the arcade version (see Figure 1).

In the ’90s, things changed. Consoles became faster, with better graphics and sound than arcade games. If you bought one of these consoles (or your parents bought one for you), you didn’t have to put a quarter into a machine every time you wanted to play the coolest game. In addition, you could play for as long as you wanted and didn’t have to travel anywhere to do so. So, the standard video arcade was able to move into the home.

It became clear that arcade games couldn’t compete with game consoles and PCs, in terms of graphics, sound, and length of play per gaming session. So, to compete with game consoles, the only thing video arcades could do was innovate at the user interface. This innovation came in the form of a variety of input devices and strategies that got players more actively involved.

In the late ’90s and early 2000s, arcade games introduced several particularly innovative user interface designs. For example, in BeachHead (see Figure 2a), the user wore a helmet-like device, resulting in a 360-degree field of regard. In games such as Football Power (see Figure 2b), players actually controlled a soccer ball with their feet. In Aliens Extermination (not shown), players used realistic gun props to interact in a first-person-shooter-style game. In Manx TT (see Figure 2c), users rode a physical motorcycle to control a virtual motorcycle. In Dance Dance Revolution (see Figure 2d), users interacted on a small dance floor. Adopting such interfaces for game consoles wasn’t cost effective, so the video arcade was able to live a little while longer.

During the '90s, VR technology appeared in video games. VR-based games helped keep arcades alive because they included more advanced interfaces employing such technologies as stereoscopic vision, head and body tracking, and 3D spatial interaction. These games appeared mainly in arcades and entertainment centers.

One of the first VR games was Dactyl Nightmare, which W Industries/Virtuality developed in the early '90s as part of a suite of VR games. Players entered a pod-like structure, put on a head-mounted display, and used a tracked joystick to interact with the virtual world. The pod protected users from walking around in the physical world. The games themselves were somewhat primitive, but the immersive experience hadn't been seen before in a video arcade game. One VR entertainment center was BattleTech, based on the BattleTech universe. The first one opened in Chicago in 1990. It aimed to provide an immersive experience where several users could play simultaneously not only at a single location but also networked across all BattleTech centers.

It's clear that game interfaces must become easier to use while maintaining the high levels of expression and control of modern console and PC games.

Although these types of video games provided more interesting interactive experiences than consoles, the cost of upkeep, lack of throughput, cost of play per gaming session, and continued improvement of game consoles led to the demise of most arcades. Arcades still exist, but they're nowhere near as popular as they were in the late '70s and '80s. Today they're often coupled with family entertainment centers, bars, bowling alleys, and other small venues. However, the video arcade game's evolution shows that as time passed, these games had to provide users with higher levels of interaction—not just using a control stick and a set of buttons—to keep them interested and coming back for more.

Input control and game complexity

If we examine the evolution of game consoles from the '70s to today, we can see profound improvements in computer graphics, sound, artificial intelligence, and storytelling. However, a closer look shows that these games have become more complex, in terms of not only story, graphics, sounds, and so on but also gameplay. As video games became more complex by giving users a greater variety of things to do and controls to master, the input devices for controlling them stayed relatively constant. If we look at PC games, the argument is clear: mouse and keyboard have been the predominant control devices since PC gaming started in the early '80s.

You could argue that console game input devices have improved and gotten more complex since the Atari 2600's simple directional joystick and button. Game controllers certainly have gotten more complex but, I would argue, not necessarily better. As new generations of game consoles emerged, the controllers simply added more buttons and joysticks to previous versions.

For example, the 1983 Nintendo Famicom used a game pad with direction buttons (to replace the directional joystick) and four buttons; the Sega Master System had a similar design. In 1994, when Nintendo introduced the first 3D game console (Nintendo 64), the controller had a directional joystick, a directional pad, and 10 buttons. Finally, the Sony Playstation introduced the DualShock controller, which has two analog joysticks, a directional pad, and 10 buttons.

As controller complexity increased, so did the complexity of a game's control scheme. These schemes allowed for more expression, at the cost of making the games difficult to learn and master. So, the interfaces became tailored to hard-core gamers, often alienating the casual player. For example, I've been playing John Madden Football since its introduction in the early '90s. With each new version, the game has gotten more realistic and improved how you control the football players and call plays. However, the game now has so many choices and controls to master that I can't remember them all, effectively limiting what I can do.

It's clear that game interfaces must become easier to use while maintaining the high levels of expression and control of modern console and PC games.

3D spatial interaction and VR in video games today

Our glance at video game history shows three trends that lead us to why 3D UIs and VR are starting to become so popular. First, once game consoles had better graphics and sound, had more interesting stories, and let users play much longer, arcade games had to give players something that they couldn't get on consoles: innovative interfaces that provided more natural means of expression. So, to compete in a market where consumers can choose from several gaming platforms, better graphics was no longer a key to staying on top. More natural gameplay, as we've seen with the Nintendo Wii's success, keeps people wanting more.

Second, as I mentioned before, more complicated video games and video game controllers gave users more expressive power but alienated casual gamers. In many cases, these games use somewhat abstract control schemes, when we consider the mappings between control mechanisms that are easy to perform naturally and spatially (running, jumping, punching, kicking, and so on) and a series of button presses. To bring back the casual gamer and improve overall gameplay, 3D spatial interaction was the natural next step.

Third, the technology to make 3D spatial interaction mainstream and not just a gimmick has arrived. Console games had previously incorporated advanced game interfaces (in the late '80s and early '90s) with devices such as the Nintento U-Force, Mattel PowerGlove, and Sega 3D glasses. However, poor technology and lack of support from game developers caused their early demise. Today, faster and cheaper sensors, faster processors that can perform complex tracking and recognition, and the need to reduce game control complexity have finally made 3D spatial interaction feasible.

Current video games employ 3D spatial interaction and VR in three ways. First, as the Sony EyeToy showed, simple vision-based tracking can let players use their bodies to control game characters.

Second, realistic 3D spatial interfaces based on active physical props—specifically, guitars and drum sets—give gamers the ability to interact as if they were in a real rock band. Guitar Hero and Rock Band are interesting examples of this type of realistic control scheme because people are willing to buy these devices for use with just one game. Ten years ago, no one would have believed that people would spend $90 to $200 on a single game with a specific controller. (This supports the possibility of more arcade-style interfaces from the '90s making their way into the home.)

Finally, and probably the most important, is Nintendo's approach with its Wii and Wiimote. The Wiimote is one of the most significant technological innovations in 3D spatial interaction for gaming. It not only acts as a gamepad but also makes games accessible to the casual gamer because it can sense 3D motion. The Wiimote's innovation lies in its overall design. It uses Bluetooth for communication, making it wireless. It senses acceleration along three axes and has an optical sensor for pointing (when a sensor bar is used). This acceleration detection gives the Wii its power in that it lets users interact with games spatially (for example, swinging a bat, tennis racket, or golf club). The device also has audio and rumble (vibration) feedback. In addition, users can easily attach different types of input devices and physical props such as tennis racket and baseball bat proxies to it.

Figure 3. A user fights enemies with sword and shield in SwordPlay. This game attempted to leverage existing 3D UI techniques.

The future

Judging by the popularity of the Wii and games such as Guitar Hero, 3D UIs have finally hit mainstream society, and gamers appear to be thirsty for more. However, these game controllers only scratch the surface of what's possible.

In addition, although these devices, especially the EyeToy and the Wii, work well, they're far from perfect, and future iterations will be required to move beyond the status quo. For example, the EyeToy wasn't designed to extract depth information with its single camera. The Wii can't detect six degrees of freedom (DOF), at least not conventionally, and can reliably handle only coarse gestural input. However, the recently announced Wii MotionPlus, an attachment to the standard Wiimote that uses three orthogonally aligned gyroscopes, will certainly improve orientation sensing.

Finally, because these devices are fairly new, many game developers don't yet understand how to fully exploit them. This trend is exemplified in that only a few Wii games make full use of the Wiimote and do it well. Most Wii games treat the controller's 3D spatial capabilities as an afterthought.

These issues present a great opportunity for 3D UI and VR researchers to determine how to best leverage existing research into games and to develop new interaction methodologies and strategies geared toward console and PC games and their constraints.

For example, in 2006, colleagues from Brown University and I wrote an article for *IEEE CG&A*'s Projects in VR department.[3] The article discussed SwordPlay (see Figure 3), a video game in which

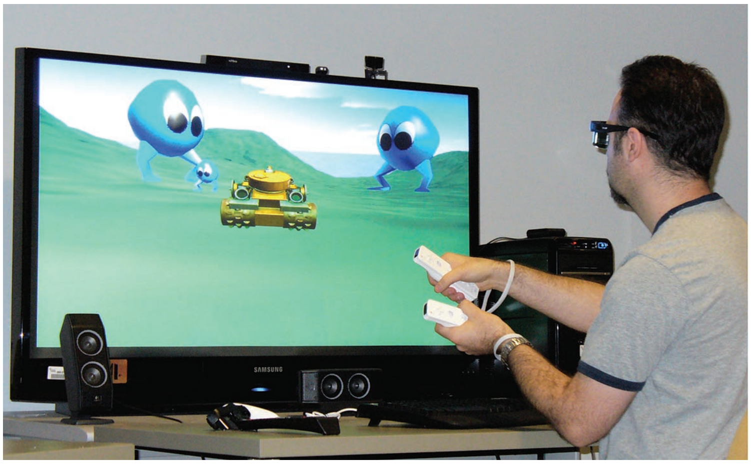


Figure 4. One of the Interactive System and User Experience Lab's 3D UI gaming stations. The user is head tracked and sees the screen in 3D stereo while controlling a virtual tank with two Wiimotes.

users fight enemies using a sword and shield, a bow and arrow, and a set of spells they can sketch with the sword. Our research aimed to leverage existing 3D UI techniques in the context of a video game. The game was played in a four-sided CAVE (Cave Automatic Virtual Environment) with 6-DOF trackers. Unfortunately, most people probably don't own a four-sided CAVE or an expensive 6-DOF tracking system. The project was successful in that it showed what might be possible in the future. However, it didn't address how to deal with the limitations of today's game input device hardware. It also didn't look at how to improve existing VR and 3D UI hardware to make it as affordable as a Wiimote or a standard TV.

To address many of these issues, my research group and I (the Interactive Systems and User Experience Lab) are exploring how to bring 3D UI techniques and concepts into mainstream video games by leveraging existing 3D UI and VR research and devising strategies and methodologies for bringing 3D spatial interaction to gamers. First, we've set up several 3D UI game stations (see Figure 4) where students and researchers can work on advancing the state of the art in game user interfaces. Each station has, at a minimum, a 50-inch 3D DLP (Digital Light Processing) high-definition TV, a set of Nintendo Wii controllers, a PC, active shutter glasses, a speaker system, and a TrackIR camera for head tracking. (Some stations have other hardware such as Novint haptic devices, webcams for exploring alternative tracking methods, and various input device tools from Phidgets and Infusion Systems for quickly prototyping new input devices.)

We chose 3D DLP TVs because 3D stereoscopic vision has been a common component of research lab VR systems for almost two decades. The ability to see in 3D stereo helps make a game more interesting and immersive. Advances in 3D stereo technology have reached the point where it's easy to have 3D stereo in the home. One such technology is Texas Instruments' 3D DLP. With 3D DLP, projection TVs provide an option where, when users wear a pair of active shutter glasses, they can view movies and play games in 3D. In addition, because these are high-definition TVs, they have excellent resolution—up to 1,920 × 1,080 pixels. The other benefit is that 3D DLP technology needs no special graphics card. So, anyone with a reasonable graphics card and the appropriate 3D content can view and play games in 3D stereo.

The basic version of one of these stations costs only $3,000. When you consider that gamers pay more than $3,000 for high-end gaming PCs, the price point for a 3D UI game station is pretty low. Additionally, because more LCD and plasma HDTVs are refreshing at 120 Hz, 3D TV will eventually not be exclusive to DLP projection.

Our software framework is based on Microsoft's XNA game development environment. It has several useful features such as a vector/matrix library, audio and speech support, Wiimote APIs and gesture recognition libraries, head tracking, physics engines, and several high-level components that make building and managing 3D scenes easier. Because 3D DLP TVs handle stereo differently from traditional stereo rendering algorithms (as I mentioned before, they let you use any reasonably powerful graphics card), the 3D UI game station's version of XNA has a modified rendering algorithm to support this form of stereo.

With our hardware and software infrastructure, we're exploring how to make rich gaming environments easier to play by finding intuitive mappings between a user's physical actions and game controls. For example, we're using head tracking in first-person-shooter games to reduce the number of button-based commands needed for crouching and peeking around corners, and we're determining how best to use 3D stereoscopic vision in games.

We're also working on Wiimote research, including determining

- what gesture sets and machine-learning algorithms work best for gesture recognition,
- how existing 3D user interaction techniques best map to Wiimote devices, and
- how to best leverage the Wiimote to support natural body-based navigation control, dance movements, and musical-instrument playing.

Finally, we're exploring ways of building a better Wiimote that uses other sensing technologies and that can be manufactured inexpensively yet supports both 3D position and orientation tracking without any restrictions on how the device should be held or pointed. Having input devices that track 6 DOF and are at the same price point as a Wiimote will go a long way to improving 3D spatial interaction and providing much more realistic and immersive experiences.

We're in the early stages of a revolution in how video games are played. Three-dimensional spatial interaction and VR concepts such as 3D stereo rendering and head tracking will play a crucial role in generations of future video games. This revolution represents a golden opportunity not only for 3D UI and VR researchers to have a "killer app" to work on but also for game designers and developers to leverage the expertise of those who have been working in 3D spatial interaction for almost two decades. Collaboration between these two separate but intertwined fields will benefit not only the parties involved but also game players, who will get to experience new and interesting ways of interacting in game worlds.

With continued research in bringing 3D spatial interaction and VR into home-based video games, we could see a time in the not-too-distant future when game consoles come with 6-DOF tracking systems and stereo cameras, gamers have dozens of input devices (perhaps with no buttons at all) for controlling specific games, and gamers play on autostereoscopic, or perhaps even holographic, TVs.

Acknowledgments

Some ideas in this article are based on part of a course I gave at CHI 08 with Doug Bowman, Ivan Poupyrev, and Ernst Kruijff.[4] I thank them for valuable discussions.

References

1. B. Sherman and A. Craig, *Understanding Virtual Reality*, Morgan Kaufmann, 2003.
2. D. Bowman et al., *3D User Interfaces: Theory and Practice*, Addison-Wesley, 2004.
3. M. Katzourin et al., "SwordPlay: Innovating Game Development through VR," *IEEE Computer Graphics and Applications*, vol. 26, no. 6, 2006, pp. 15–19.
4. E. Kruijff et al., "3D User Interfaces: From Lab to Living Room," Course 17, presented at 2008 SIGCHI Conf. Human Factors in Computing Systems (CHI 08), 2008, www.3dui.org.

Contact author Joseph J. LaViola Jr. at jjl@eecs.ucf.edu.