Immersive 3DUI on One Dollar a Day

Aryabrata Basu\(^1\), Christian Sautpe\(^2\), Eric Refour\(^3\) \hspace{1cm} Andrew Raij\(^4\) \hspace{1cm} Kyle Johnsen\(^5\)

University of Georgia \hspace{1cm} University of South Florida \hspace{1cm} University of Georgia

ABSTRACT

A convergence between consumer electronics and virtual reality is occurring. We present an immersive head-mounted-display-based, wearable 3D user interface that is inexpensive (less than $900 USD), robust (sourceless tracking), and portable (lightweight and untethered). While the current display has known deficiencies, the user tracking quality is within the constraints of many existing applications, while the portability and cost offers opportunities for innovative applications that are not currently feasible.

KEYWORDS: virtual reality, ubiquitous computing, wearable computing, gaming

INDEX TERMS: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

1 INTRODUCTION

In 1991, Randolph Pausch published a paper entitled “Virtual Reality on Five Dollars a Day” describing the design of a $5000 (USD) virtual reality (VR) interface; which, amortized over three years, was about five dollars a day [1]. To build such a low-cost system, off-the-shelf hardware was combined with creative engineering to form a makeshift, but complete, immersive 3D user interface (3DUI). The cost was a stark contrast to expensive commercial systems available during that time. More important than the price, however, was the possibility of immersive technology being available to vast numbers of designers and end-users, enabling immense creative efforts and sparking a renaissance of VR. Two decades later, this possibility is now rapidly becoming a reality. In this paper, we present an evolution of this idea, a complete, immersive 3DUI for one dollar a day.

Enabling this evolution are low-cost consumer electronics devices that are mass-produced for entertainment purposes, yet are essentially the same technologies once reserved for VR applications, and often have as-good or better performance than existing “professional” devices [2-4]. These consumer devices are thus viable alternatives for 3DUI designers, and have indeed been particularly popular for prototyping new systems [5, 6, 7].

Building upon this idea, we have previously reported on the design of a mobile-VR system for immersive collaborative virtual environments [8]. The goal was to design a low cost system that would allow a user to enter a shared virtual space from anywhere, with the immediacy of a phone-call. Our approach combined a networked smart phone device with its embedded motion sensors and a connected head-mounted-display (HMD). The effect was to produce a minimal virtual reality system that could be used within seconds of the user’s desire to enter a shared virtual space.

The primary limitation of the previous design was that only

\(^1\)basuarya@uga.edu \(^2\)christiansaupe88@gmail.com \(^3\)erefour@uga.edu \(^4\)raij@usf.edu \(^5\)kjohnsen@uga.edu

IEEE Symposium on 3D User Interfaces 2012
4-5 March, Orange County, CA, USA
978-1-4673-1205-9/12/$31.00 ©2012 IEEE
with portable video devices (e.g. Vuzix, EMagin), and motion-
controllers have become increasingly popular for gaming (e.g.
Nintendo Wii Remote, Microsoft Kinect, Sony Playstation Move).

Furthermore, smart phones and similar mobile devices are available that have powerful processing, graphics, and display
capabilities (e.g. Apple iPhone & iPod Touch). While the
performance may lag behind that of traditional VR systems, they are
rapidly improving in performance with GHz multi-core CPUs
dedicated GPUs. In fact, Olsen et al demonstrated the use of
these devices as a stereoscopic HMD with custom optics [14].

Our contribution is to merge related work in portable, low-cost VR, providing a high degree of immersiveness and interactivity at
a price that allows for ubiquitous deployment of such systems.

3 SYSTEM

According to Pausch, the foundation of an immersive system
consists of a tracked HMD supporting immersive viewing, a hand
held or worn tracked device supporting immersive interaction, and
a computer to integrate tracking and render the virtual world [1].

Display: A lightweight HMD (The Vuzix Wrap 920) is used for
the display. The HMD has two independent 640 x 480 24-bit color
liquid crystal displays (LCDs), and supports stereoscopic
rendering via a side-by-side display format (in stereoscopic mode,
each rendered view is 320 x 480 and is interpolated to 640 x 480).
The aspect ratio is 4:3, with a 30deg diagonal field-of-view. While
it has a low field-of-view and resolution, the lightweight (110g),
battery powered (2 AA for 2 hours running time) design makes the
Vuzix HMD well suited for a mobile display system. The
Vuzix HMD was modified slightly by replacing the sunglasses-
style mounting with an elastic band in order to more securely and
comfortably bind the display to the wearer’s head and provide
more convenient mounting of tracking devices.

Tracking and Interaction: Two tracking and interaction
devices are used. The first is the sensor system built into the iPod
Touch 4g. This device has a 3-axis accelerometer (16-bit, -2.4g, to
2.4g), and a 3-axis gyroscope (16-bit, -2000deg/s to 2000deg/s).
A filter is used to integrate accelerometer and gyroscope readings
and produce an orientation that is correct about the axis of gravity.
The Apple iOS Core Motion library performs this filtering and
provides a unit-quaternion representing the orientation at 100Hz.

The second tracking system is the Razer Hydra. The Razer
Hydra is the first mass-produced magnetic tracking system intended for
the video game market. It provides 250Hz 3DOF position and orientation of two wired hand-held controllers, each of which has a joystick and 8 buttons. It is small (see Figure 2),
lightweight (800g), and powered through the USB system. The
position and orientation computations are performed on device,
and can be obtained through a free SDK from Sixense (developers of
the Razer Hydra), or through a virtual reality peripheral network (VRPN) server developed by Ryan Plavik
(https://github.com/rplavik/razer-hydra-hid-protocol). We learned that the magnetic tracking sensor could be removed from the
Razer Hydra controller, making it much smaller and lighter at a
loss of the buttons and joystick (Figure 2).

Computing: Two mobile computers are used. The first is the
BeagleBoard XM single board computer. The BeagleBoard XM
includes a Texas Instruments DM3730 System-on-Chip (1GHz
Arm Cortex A8 processor, 512MB memory, and PowerVR
SGX530 graphics chip). It has interfaces for 4 USB devices,
audio, S-Video, and HDMI. As the vast majority of consumer-
level interaction devices (e.g. the Razer Hydra) have USB
interfaces, the availability of powered USB connections was an
important consideration. The BeagleBoard XM can run the Linux,

Table 1. Retail cost of hardware components

<table>
<thead>
<tr>
<th>Item</th>
<th>Retail Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vuzix Wrap 920</td>
<td>$299.99</td>
</tr>
<tr>
<td>Razer Hydra</td>
<td>$139.99</td>
</tr>
<tr>
<td>Apple iPod Touch 4g (8GB)</td>
<td>$199.99</td>
</tr>
<tr>
<td>Beagleboard XM</td>
<td>$149.00</td>
</tr>
<tr>
<td>Beaglejuice 4500mAh Battery</td>
<td>$88.73</td>
</tr>
</tbody>
</table>

Total $873.70

Figure 2. The Razer Hydra and its decomposition to remove the circuit board containing the magnetic sensing coils

Android, or Windows CE operating systems. Linux was used in this work (Ubuntu 9.10).

The second computer is the iPod Touch 4g. It contains an
Apple A4 System-on-Chip (800MHz Arm Cortex A8, PowerVR
SGX540 GPU, 256MB memory), built-in WiFi and Bluetooth
networking, and has a proprietary connector that can be used to
attach external devices including displays. The iPod Touch was
primarily chosen because of its wide customer base (millions of
devices), software distribution mechanism (Apple AppStore),
impressive embedded sensors, and because it connects directly to
the Vuzix HMD. Use of other smart-phones and platforms is
possible, provided they can attach to the HMD. For example, the
Beagleboard’s s-video output could also drive the HMD.

Design: As shown in Figure 1, the Razer Hydra source is
mounted to the back of the belt. The iPod Touch is clipped to one
side and connected to the HMD controller box. The Beagleboard
is mounted to the other side and connected through USB to the
Razer Hydra and a WiFi adapter. One of the two Razer Hydra
pose sensors is separated from its controller body and attached to
the HMD. The user holds the other controller. This design greatly
increases immersion with respect to the previous design. The
Razer Hydra provides robust 6-DOF pose tracking for the user’s
head and hand relative to the hips of the user. The orientation of
the hips is tracked by the iPod Touch inertial sensors, and thus no
functionality is lost. Furthermore, the head and hand tracking now
have the same frame of reference (the hips) and tracking
performance characteristics, which is important for maintaining
consistency (although we introduce a latency discrepancy when
the hips are moving, as discussed in section 4).

Mounting the iPod Touch at the hip, instead of the head, allows
the use of several common locomotion metaphors. In addition to
moving by pointing or with the joystick, the accelerometer can be
used to detect motion. While double integration of accelerometer
is theoretically possible, numerical error accumulation makes this
infeasible. However, footfalls can be detected reliably (see
Section 4), allowing locomotion by walking in place or real
walking (provided space is available). All of these techniques
offer only relative motion. Absolute motion currently requires an
external system, e.g. GPS or fiducial tracking markers.

An ad-hoc WiFi network connects the Beagleboard and iPod
Touch. Data from the Razer Hydra is read by software on the
Beagleboard and transmitted using the VRPN library to the iPod.
Table 2. Average frames per second for test scene by device and number of virtual human avatars (11200 polygons each) in the scene. *The hardware is limited to 60 frames per second.

<table>
<thead>
<tr>
<th>Avatars</th>
<th>iPod Touch 4g</th>
<th>iPhone 4s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60*</td>
<td>60*</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>60*</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>48</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>26</td>
</tr>
</tbody>
</table>

The iPod then converts the incoming pose data for the head and hand into its own reference frame (as obtained from its inertial sensors).

**Cost:** Table 1 shows the cost of each component (as of this writing), ignoring cabling and attachment materials. The total cost for the system is $873.70 (USD). If amortized over three years (as in Pausch’s work), the cost is $0.80 (USD) per day. For simplicity, and the fact that some application-specific modification will be needed, we round up to $1 per day. This is one fifth of the cost of Pausch’s system (lower if inflation is taken into account), while offering far higher performance and mobility.

### 4 Performance Characteristics

To test the performance of the user tracking and devices we conducted comparisons with a 5-camera NaturalPoint Optitrack optical tracking system. This tracker provides 6-DOF pose measurements of rigid constellations of reflective infrared spheres at a 100Hz update rate. Within the 3m x 3m x 3m tracking volume it has excellent accuracy (< 1cm), resolution (< 1mm), and latency (10ms). All data was collected on the iPod Touch, which sampled at 60Hz (its maximum frame rate for 3D applications). No filtering was performed on the data. All objects (the iPod touch, Razer Hydra source, head sensor, and wand sensor) were tracked by the Optitrack system.

For the first test, we compared the orientation measured by the iPod Touch inertial sensors to that measured by the Optitrack system. We collected a 30 second sample of rapidly swinging orientation motions from both the iPod Touch and an Optitrack system rigidly attached to the iPod (timestamps were recorded on the iPod as data arrived from the Optitrack system over VRPN). We converted the quaternion obtained from each tracking system to an axis-angle notation and used the angle as the comparison metric. We noticed a substantial latency from the iPod Touch orientation sensor with respect to the Optitrack sensor. By time shifting the Optitrack data until the error between the two sensors was minimized, we determined this latency to be 41ms (this is in addition to Optitrack and network latencies). The average absolute error between the two measurements was approximately 4 degrees.

For the second test, we compared the position accuracy of our hybrid inertial-magnetic tracking system to the NaturalPoint Optitrack system. The average absolute errors were 24.6mm, 13.4mm, and 13.5mm for the x (left), y (up), and z (out) axes respectively. Latency, in this case, was difficult to determine, as it is a combination of the two tracking systems employed; however, at times there appeared to be no latency between the two systems. This can be explained by the combination of the systems. The Razer Hydra has a marketed 4ms latency. This is less than the 10ms latency of the NaturalPoint Optitrack system. Our measured latency of the iPod touch inertial sensor is at least 41ms. Thus, it depends on which one is currently varying as to what latency will be perceived. The two most important components, head and hand tracking are both measured with the low-latency Razer Hydra.

For the last test, we measured the rendering performance of the system. A simple test scene was composed in the Unity 3D game engine. Unity 3D was chosen for convenience, and could be replaced by a free alternative such as Ogre 3D (as was used in our earlier work). The scene consisted of a number of articulated virtual humans and was indicative of the environments that we envisioned the system would be used for: social, collaborative environments. The virtual humans were obtained from www.evolver.com and were each 11200 triangles. To test, we varied the number of virtual humans visible in the window. The environment was rendered in a side-by-side viewing configuration for the left and right eyes (as needed by the HMD). For comparison, we also measure frame rates for the more powerful iPhone 4s (Apple A5, 800MHz Dual Core). The iPod Touch 4g performance was about half that of the iPhone 4s (Table 2). Interactive frame rates were achieved in all cases, although the eight-avatar case for the iPod Touch 4g was only marginally acceptable at 13 frames per second.

### 5 Discussion

The performance experiments addressed areas of concern related to the inexpensive approach. First, we showed that the iPod Touch 4g inertial sensor was capable of accurately tracking the orientation of the user. The errors were low (approximately 4 degrees) and only accumulated about the axis of gravity. However, it did have a high latency (approximately 41ms). This was a concern with the previous head-mounted design, as rendering was directly coupled to the tracking system. In the new approach, however, latency in the inertial tracker is indirectly coupled to the magnetic tracker, only affecting the view when the hips are rotated. Thus, if hip rotation is infrequent for an application, the latency will not be a large source of concern.

The purpose of the magnetic tracker was to enable body-centric position tracking. In this regard, the performance of the system was exceptional. Magnetic tracking is well suited for body centric
interaction, particularly when the magnetic source is mounted to the body. This alleviates some of the primary causes of error associated with magnetic tracking, namely distance from the source (the source travels with the user) and magnetic field distortions (the body does not distort magnetic fields). In fact, we noted during the experiments that the Optitrack system frequently lost track, and occasionally flipped orientations, making it the less robust of the two tracking systems for body-centric tracking. The Razer Hydra, in particular, is a high quality product for its price, and the ability to remove just the magnetic field sensor from the body of the wand makes it even more flexible.

Lastly, we note that the rendering performance of the iPod Touch 4g was adequate for many VR systems. Also, we found that smart phone performance is exponentially increasing with each generation. It is likely that rendering performance will not be a major issue for future generations of this concept.

5.1 Limitations

There are some limitations to the tracking approach. First, the system does not support crouching or climbing, because it technically cannot detect the height of the user’s hips above the ground. We could incorporate additional magnetic sensors on the feet and torso. This would enable crouching to be detected, and would improve locomotion.

Another related limitation is that finger tracking is not supported. While inexpensive data-gloves were examined for this system (e.g., P5 Glove and the Nintendo PowerGlove), these were not of sufficient quality to incorporate. Low cost optical tracking approaches show promise in this area.

The largest remaining concern is the lack of an inexpensive large field-of-view head-mounted display. The Vuze VR920 has reasonable visual quality, and its lightweight design makes it comfortable to wear for extended time, but its low field of view makes achieving a sense of presence difficult. For this reason, we did not try to block out the real world (Figure 1). Thus, current applications will likely be oriented towards social gatherings and entertainment rather than those relying on high presence such as exposure therapy.

5.2 Research Questions and Challenges

The nature of an inexpensive, portable, untethered VR system poses intriguing new research questions and challenges, specifically related to the idea that VR or 3DUI experiences are likely to occur in uncontrolled environments. Given that VR is concerned with the virtual world, the overarching question is “what do we do with the real world?”

Effectiveness & Distraction: Can we build VE s or 3DUIs that are effective, despite the often-unpredictable distortions present in the real world (e.g., a knock on the door, a glaring ambulance driving past, or drops of rain when outdoors)? Furthermore, how do we evaluate the effectiveness of VE s or 3DUIs in uncontrolled environments? Traditional measures of presence may not be appropriate for VE s where distortion is the norm.

Hiding Reality: Should portable VR or 3DUI systems like the one presented here block out the outside world? For example, redirected walking [15] and other techniques based on perceptual illusions could be used to minimize the chance the user collides with a real wall. Similarly, if it starts to rain in the real world, the system could generate rain in the virtual world to minimize distraction from the unexpected external stimuli.

Leveraging Reality: Alternatively, could characteristics of the real world be used to improve theVE or 3DUI? For example, if a map of the user’s external environment were available, one could automatically align the VE with the real world (e.g., align real and virtual walls) to provide passive haptic feedback. Similarly, if the virtual experience takes place in a rainy outdoor environment, the user could enter the VE while standing outdoors in the rain.

6 Conclusions

With the low cost system described in this paper, the monetary barrier to entry in immersive VR is all but eliminated. For less than the price of a mid-range television (particularly if the user already has a smart phone), a user may interact in immersive virtual worlds. It is possible that in the near future, VR will become ubiquitous, but for that to occur, mass-appeal applications are needed. Our hope is that the approach presented in this paper could serve as a catalyst for creating such applications.

Our future work with this system is currently targeted towards large-scale collaborative interactions for “second-class” applications that cannot afford large-scale virtual reality installations, such as education and entertainment. With the current design it is possible to deploy hundreds of immersive systems in places that were once never thought viable.

References:


