Abstract

Virtual Reality through head-mounted displays (HMD) has great potential for further development in various applications. There are still several problems that must be solved in order to make a product such as this, mainstream. Besides creating a lacking immersive virtual reality experience, fundamental issues such as time lag and motion sickness need to be resolved. This can be done with more dedicated research into the efficiency and accuracy of sensor hardware and the sensor fusion algorithms that utilize the data. Putting more time and money into developing an efficient approach to multi-sensor data fusion will not only lower the costs of hardware, but also improve the users immersive experience and make it a true virtual reality. This paper analyzes the history, technical underpinnings, a demonstration, and future trends of immersive first person virtual reality.

1 INTRODUCTION

Immersive First Person Virtual Reality is rapidly developing in today’s entertainment industry. The revival of this field will bring with it a wide
array of helpful applications for the world to use. With a relatively short lifespan, virtual reality has had a difficult start but is sure to see improvement in the near future.

2 HISTORY

The history of head-mounted displays dates back to 1965 with Ivan Sutherland, an Associate Professor of Electrical Engineering at Harvard University, where he established the principal notion of virtual reality with The Ultimate Display. This idea was to make a (virtual) world in the window look real, sound real, feel real, and respond realistically to the viewers actions. [15] It involved using a display connected to a digital computer in order to gain familiarity with concepts not realizable in the physical world. Sutherland furthered his research in 1968 by developing the first known HMD called The Sword of Damocles. This prototype provided the user with a perspective image, which changed as he/she moved. In 1984, Jaron Lanier, known as the man who coined the term Virtual Reality, founded VPL Research, a company who went on to create products such as The DataGlove, The EyePhone, and The AudioSphere. VPL filed for bankruptcy in 1990 and later in 1999, all of its patents were bought out by Sun Microsystems, a company that sold computer components, computer software, and information technology services such as the Java Programming Language. In 1995, a research team at the University of North Carolina developed a video see-through HMD. This display gave the users a view of the real world through one or more video cameras mounted on the display. Synthetic imagery was combined with the images captured through the cameras. The combined images were sent to the HMD. At the beginning of the 21st century, personal electronics had become extremely relevant, with private companies introducing smaller and smaller multi-purpose devices every year, VR HMDs partially followed this trend but was still not very consumer friendly. Many major companies such as Sony and Zeiss developed consumer HMDs but were, and still are, extremely expensive. Most of these served less of an immersive virtual reality experience, and more of a 3D movie experience. Though this is only a brief look at the main moments in virtual reality history, a lot can be told about the prior status quo, the pressing need for the development of this technology, the catalysts that were involved, the introduction phase, and the new status quo.

2.1 Prior Status Quo

The prior status quo to virtual reality technology was of mere imagination and demonstration through art. The first of it’s kind that could be close to being called virtual reality was a mural by Baldassare Peruzzi. This mural was a room that encompassed the viewer in a painted alternate reality. [7] This adaptation of virtual reality was formed in the 1860’s and remained the only prominent idea until the 1960’s when Morton Heilig’s ”Sensorama” was developed. It encapsulated the increasing demands for ways and means to teach and train individuals without actually subjecting them to possible hazards of particular situations. [10]

2.2 Why was it needed?

We have yet to see a dire necessity for virtual reality hardware and applications. That’s not to say that there hasn’t been helpful VR models out there that have improved human life. Development and research into virtual reality saw a
glimmer of hope when the idea of training simulators became prominent. The idea behind virtual reality simulators was to keep the students (or the clients) out of harms way for as long as possible until training was sufficient enough for technical operations to be close to perfect. Examples of training include surgical, mechanical, automotive, and geographically dangerous operations. [20]

2.3 Catalysts for Development

A few of the many catalysts involved in furthering the development of virtual reality hardware and applications was improving the overall safety of clients in many different environments. Virtual reality applications were needed and targeted toward training professionals for complex or risky procedures that have become a lot more common and prove to show better results compared to brochures, manuals, or videos. Some of these training programs were involved in different types of invasive surgeries, dangerous mining operations, flying many different types of airplanes in a wide array of dangerous weather conditions, or driving a car in different types of conditions.

2.4 Introduction Phase

The frontier of VR digital simulation began in the early 1950s initiated by the University of Pennsylvania. This piece of hardware was named UDOFT (Universal Digital Operational Flight Trainer), which was run by two parallel processors. By the 1970’s, general purpose computers had improved to the extent that they could be considered for use with flight simulator applications. [17]

2.5 New Status Quo

The state-of-the-field of virtual reality displays is currently in a research and development phase, with only a few major companies such as Sony, Zeiss, Sensics, and Oculus making noticeable headway. It still faces many difficult research problems involving many disciplines. Thus, realistically, major progress requires decades rather than months. The expense of researching virtual reality is extremely high and unfortunately the support for this type of technology is relatively lacking. The major consumer products that are currently being used are for gaming and cinema. This includes the Oculus Rift for immersive gaming, and several cinematic headsets such as Sony’s Wearable HDTV device.

3 TECHNICAL ANALYSIS

In order for virtual reality applications to be truly immersive, the experience needs to be flawless. The hardware needs to be light enough to fit comfortable on the head for long periods of time. The displays need to be advanced enough to surround the user in a seemingly realistic virtual environment. The visual applications must be smooth enough to eliminate motion blur, time lag, screen tearing, which as a result could cause sickness to the user. In virtual reality systems the head-mounted display has many jobs. The most important of these jobs is "to know the relative position and orientation, known as the pose, between the head and the virtual environment surrounding the user.” [11] The pose is also known as the 3D vector of position and orientation. To determine the pose of an object with respect to the user’s head, tracking sensors are necessary. Sensor technologies that have been used in the past
include mechanical, magnetic, acoustic, and optical. Estimations are made by these sensors, if the estimations are done incorrectly the users view within the virtual world will be inaccurate. "Registration inaccuracy is one of the most important problems limiting virtual reality applications today." [11] Fixing these inaccuracies involves every technical aspect of the head-mounted display system. Sensor hardware must be extremely quick in providing accurate raw data to the application. The raw data from each individual sensor only gives us a limited view of orientation at a certain point in time. Combining all of the sensors data, analyzing it, and outputting useful information to the application offers a much more streamlined performance for the user. There are several major multi-sensor data fusion algorithms that are used for head-mounted displays, the Central Limit Theorem, Kalman Filter, Bayesian Networks, and the Dempster-Shafer algorithm. [21] All of which are able to serve different purposes, but the main sledgehammer for this type of work is the Kalman Filter algorithm, but as it is well known in software development, there is no one way to program an application to perform a task, but there can be a current best way. This leads us to some of the issues and questions about the technical aspects of sensor hardware and algorithms. There is still not enough money, research, or development going into this field to tell which approach is the most efficient and cost effective.

3.1 Technical Aspects Head-Mounted Displays

Head-mounted displays have several technical aspects to them, surrounding hardware, data sensors, and software algorithms combined with those sensors. One of the main challenges with head-mounted displays has always been the form factor. In the beginning of development for the first head-mounted display, the device needed to be tethered to the ceiling of the facility that it was located, in order for the user to have it placed on (or should we say float above) their head. With the help of 20 plus years in research, these devices have turned into "lightweight, (somewhat) mobile, ergonomic optomechanical headsets that can fit comfortably on the users head." [12]

The four major fundamentals of a head-mounted display starts with a microdisplay. Microdisplays within HMDs need to provide a fairly high resolution in a relatively small area. The challenge with this is keeping the screen small enough to fit comfortably on the person, but at the same time having the display be large and advance enough to provide a realistic virtual reality experience. The second fundamental aspect of immersive HMDs are the optics, i.e. the designs that block the direct real-world view, and replaces it with a virtual environment. "Ideally, immersive HMDs target to match the image characteristics of the human-visual system. Because it is extremely
challenging to design immersive displays to match both the field of view (FOV) and the visual acuity of human eyes, tradeoffs are often made." [12] Many different HMDs used today are utilizing several different styles of lenses. Some lenses allow for the use of glasses or contacts, other lenses offer a wider FOV with greater warping, and some a smaller FOV with less warping. The third technical aspect of HMDs are the hardware sensors. There are several sensors that can be used to determine the pose of the users head within a virtual environment. The majority of sensors used, but are not limited to, are gyroscopes, compasses, accelerometers, linear accelerometers, and gravity sensors. The gyroscope measures the rate or rotation in rad/s around a device’s x, y, and z axis. The compass measures the earth’s magnetic field and outputs a value representing the current heading. The heading is calculated to a certain degree and is returned as a number from 0 to 359. The accelerometer measures the acceleration applied to the device, including the force of gravity. The linear accelerometer provides you with a three-dimensional vector representing acceleration along each device axis, excluding gravity. Finally, the gravity sensor provides a three dimensional vector indicating the direction and magnitude of gravity. These sensors are the key components to allowing applications to function at a high level of efficiency. [3] Now that we know the technical hardware of a head-mounted display, how does it handle the data that is gathered by the many sensors inside? Sensors gather a substantial amount very important data in a very small amount of time. This is due to the fact that the pose of the user in the virtual world needs to be computed before the user notices a difference in the real world. If this computation is late, it leads to something called time lag, the amount of time the computation takes to adjust the estimation of a users real-world pose compared to their virtual pose. "One of the most widely acknowledged shortcomings of current HMD systems is the time delay from the user’s head motion to the display output. Time delay causes the uncomfortable mismatching between the sense of motion and the visual input for
the user, in the worst cases causing a sort of motion sickness (virtual sickness).”[16] "To prevent motion sickness, rendering latencies lower than 10 ms are necessary.”[18] If there are too many sensors sending an overwhelming amount of data to the CPU, it can become late with its pose estimations. So why don’t we just use one sensor worth of data to compute the users pose, won’t it be faster and more accurate with less sensors? The answer is no, as provided in the information above, each sensor has a specific job, each additional sensor to the head-mounted display adds to the accuracy of the pose but adds the ambiguity of more data that needs to be computed into usable information by the application. "If the data from multiple sensors are complementary, then the resulting pose can be much more accurate than that from each sensor used alone. In order to do this, a mathematical analysis is required of all errors and uncertainties associated with the measurements and derived poses.”[11] This is done by multi-sensor data fusion algorithms, our fourth fundamental aspect, or in short sensor fusion. These algorithms are "the process of combining observations from a number of different sensors to provide a robust and complete description of the environment and process of interest. Data fusion systems are often complex combinations of sensor devices, processing, and fusion algorithms.”[13] Out of the many tools available for data fusion, there are a few essentials that are reasonably well established in the field. These include the Kalman Filter, and Bayes’ Rule. However, these tools still remain in a highly researched and developmental cycle for use in realistic applications.

3.2 Sensor Fusion Algorithm: The Kalman Filter

![Kalman Filter Equation]

Sensor fusion involves a wide spectrum of areas, ranging from hardware for sensors and data acquisition, through analog and digital processing of the data, up to symbolic analysis all within a theoretical framework that solves some class of problem. Multiple sensors in a control system can be used to provide more information, robustness, and complementary information. [8] As many resources have told me again and again, it’s nearly impossible to grasp the full meaning of Kalman Filter by starting from definitions and complicated equations. The following is what I am able to start with, which seems to be the easiest to understand.

Without getting into too many details yet, in my opinion the Kalman Filter was explained best by this definition: "The Kalman filter finds the most optimum averaging factor for each consequent state. Also somehow remembers a little bit about the past states.”[6] So what exactly does that mean? Kalman filters are a way of filtering the errors out of your sensors, and figuring out from sample to sample what sensors you can trust and what sensors you cant. It filters out certain unwanted noise in the sensor measurement.

In the figure above you can see an example of noise created by gps sensor data. The red line is the data that we’re getting from our actual sensor, where the dark blue line is the actual path.
that the vehicle took. What we have is the basic shape of the path with a bunch of noise randomly surrounding it. If we were to just average out the noise, you will not gain useful information due to slowing down sensor readings. Instead of just taking readings, you now have to take time to average them.

![Figure 5: Kalman Filter Noise Estimations](image1)

In the figure below you can see actual stationary gps data. The center of the circle is the point that it is on the earth. The points are the distribution of errors, where the center is where you really are. If we plot this data on a histogram, we would get low values, a curve to a peak in the center, a curve back down, and then low values for the remaining of the x-axis. The peak will be the most accurate data, where the lower floor will be the least accurate data. What we want to do is remove the lower floor of data, or filter out all of the errors. This is easy when you’re stationary, but for the most part, whether you’re in a vehicle or inside a head-mounted display, you’re constantly moving. Estimating errors on the move can be a little more difficult. We have to estimate given our sensor data, each sensor has a certain amount of accuracy and a certain amount of noise associated with it. We can use the equations of motion to estimate where the head has moved, and then use the sensor reading to apply if those estimates make sense. We will updates the estimates of the error and the accuracy of each sensor from step to step and then start over again, that’s what Kalman Filter does. [9] The Kalman Filter equation includes the prior knowledge of our state, increment the time step, make an estimate on where we think the sensor is going to be, we compare our prediction to the estimate of what our sensors are telling us, then we use that data to update the estimate of error from the sensors. [9]

![Figure 6: Stationary GPS Accuracy](image2)

3.3 Issues and Open Research Questions

As described above, there are several issues and open research questions regarding the technical aspects of virtual reality application and head-mounted displays. The first of which is the time delay. "A VR/MR system is composed of several components serially connected. Each component also has its independent transaction time cycle or time quantization, and this unsynchronism adds the uncertain amount of delay." [16] A time delay is a result of the following happening.

1. The sensing device measures the head loca-
2. The device then transfers this data to the computer.

3. The data is received by the computer, via device driver, and reaches the application process.

4. The interaction and simulation calculation are performed in the user application part.

5. The image is then stored in the VRAM.

6. The content of the VRAM is transferred.

7. The display shows the graphics.

"Within experience the basic delay time observed ranges from 80 milliseconds to 120 milliseconds." [16] That is not a lot of time to work with, thus shows the importance of the efficiency of sensor fusion algorithms.

Some new questions have arrived in the form of Google Glass, a new kind of head-mounted display.

Figure 7: Google Glass

Head-Mounted Projected Displays (HMPD) are a new shift in the paradigm of HMDs. "Their design allows the replacement of compound eye-pieces with projection optics combined with phase conjugate material, known as HMPDs. They consist of miniature projection lens(es), and micro-displays which shows a non-distorting retroreflective sheeting material within the environment." [12] These types of HMDs have a whole different potential of applications. Instead of being just limited to the virtual environment within the microdisplay of the HMD, HMPDs have the opportunity to work with a virtual projection overlapping on top of the users real world view.

Multisensor fusion techniques have been applied to a wide range of problem domains. Including: mobile robots, autonomous systems, object recognition, navigation, target tracking, etc." [8] With every year of research and development, limitations are declining and new domains are arising.

However, one very important new area that we have not covered is the ability of multisensory fusion systems to learn during execution. These are very exciting times, and we believe that major strides will be made in all these areas in the next few years. [8] This type of learning ability from an HMD branches out into artificial intelligence which is quite the controversial topic on its own.

4 DEMONSTRATION

For my demonstration and prototype in order to further the ideas within my research of virtual reality, I utilized several things. Due to the expensive nature and availability of immersive first-person head-mounted displays, I was forced to replicate the experience with my personal tablet computer. HMDs have several sensors, all varying from headset to headset, but conveniently enough much of the same sensors are used in
personal computers such as phones and tablets. In order to improve my research on this field, I wanted to look more deeply into sensor fusion, and really figure out how the combination of these sensors and algorithms affect the user experience within an application. I picked two sensors from my tablet to analyze. The accelerometer giving acceleration data, and the gyroscope giving orientation data. The raw data was taken from the tablet via a program called Sensor Simulator. Sensor simulator takes a snapshot of the sensors in your hardware every second over a predetermined time frame. I wanted to recreate a moment that simulated motion lag and defective pose measurements from a user who just shook his head a lot to purposely destruct sensor data. By starting the Sensor Simulator and holding my tablet at a default orientation without moving, then shaking the tablet rigorously over the course of 18 seconds, the raw sensor data that was gathered showed that a the default orientation was unable to be reached and the end of the simulation. As you can see in the figures below, I have graphed orientation and acceleration over time to get a better understanding of what is happening to the sensors when we attempt to make it ineffective within an application. As acceleration increased rapidly, orientation was completely lost and took several seconds to regain an idea of the original pose of the user, but not without error.

5 FUTURE TRENDS

Since the beginning of the field of virtual reality, most applications and hardware revolved around the use of the head-mounted display, and still do. This type of VR experience is also commonly used with the inclusion of additional sensors such as cameras or limb movement sensors. As the years continue on, we see more and more variations of the HMD and motion sensors. This has the possibility to change in the next 5 years in the industry. Virtual reality applications and head-mounted displays are receiving much more attention than they ever have in the past. The reason, virtual reality hardware is becoming cheaper and will soon be in production as consumer friendly devices. There is still much development to be done with HMDs and virtual reality applications, but with the recent acquisitions of major producers of virtual reality hardware, they will come very quickly. The most prominent issues that will be addressed during the next half decade will be cost, size, user mobility, single and multi-user interactivity, and enhanced displays. [14]
This paper will be predicting what the state of the field for virtual reality applications and head-mounted displays will be like in 3-5 years.

### 5.1 VR HMD’s State of the Field in 3-5 Years

Wearable VR has already been well established and is on its way to becoming a mainstream entertainment, and hopefully educational, tool of the future. There are several areas where this field will be improving in the next 5 years. These areas include improvement to head-mounted displays, virtual reality applications, sensor algorithms, and instantiation of augmented reality devices and applications. Head mounted displays typically have limited resolution, can create many different health issues, and user disorientation. Research in the VR HMD field will hope to solve all of these issues. Leading developers such as Oculus Rift and Sony’s Project Morpheus will have the best opportunity to do so. These future HMDs are meant to create an unmatched entertainment experience. Improvements will include 1080p resolutions for each individual stereo display and a 90 degree field of view. Future HMDs will include 3D audio technology that re-creates stereoscopic sounds in all directions and changes in real-time depending on your head orientation. [22] Improvements to multiple sensors will be made to reduce, or completely remove, latency/display lag issues, which in turn will eliminate motion sickness for most users. Displays will be improved dramatically in quality but may also increase the overall weight of the HMD. The goal for most developers will be to increase comfortability of HMDs in order to assure an immersive virtual reality experience. We can divide virtual reality applications into two main avenues, entertainment, and education. The future of virtual reality entertainment will be through interactive movie and gaming environments. We’ve already seen 3D movies for several years now, but it could be possible to immerse yourself into the movie environment completely with the use of head-mounted displays. Virtual reality gaming is underway to being incorporated into the family living room by the end of this 2014 summer. Users will soon be able to purchase a relatively cheap consumer HMD to play and interact with their favorite video games in an immersive virtual environment. The real value of upcoming research and development in this field is with educational applications. There are 3 aspects of learning that come into play for...
any one individual: visual, auditory, and kine-
thetic. Visual learners prefer seeing, auditory
learners prefer hearing, and kinesthetic learners
prefer moving, touching, and doing. The future
of virtual reality applications in education would
have the ability to combine all 3 aspects of learn-
ing on a more personalized level for each user.
Enhancements to certain aspects can be made
in order to cater to ones preference or to find
out which works more efficiently. Future mo-
tion sensing devices coupled with HMDs have
the potential to provide a learning machine that
can be utilized not only in schools but at home
as well. All in all, students will be more engaged
in the learning process. [4] In some cases, helpful
virtual reality applications have been used to del-
iver exposure therapy to assess PTSD and pro-
vide stress resilience training for those affected
by the disorder. [2] Prediction algorithms that
use several sensors, including cameras, will ro-
tate the virtual world intuitively in real-time.
[22] There are several sensor fusion algorithms
that have been standardized in the past decade.
These are less likely to change due to theoret-
ical and mathematical limitations. Augmented
reality differs from virtual reality in the types of
displays and applications that are used. The dis-
plays include head-attached displays, retinal dis-
plays, head-mounted projectors, hand-held dis-
plays, and spatial displays. Application use is ex-
panded to altering the current reality around the
user, not just a virtual environment. Time criti-
cal rendering algorithms, methods for predictive
tracking, and eye tracking technologies have the
potential to yield registration improvements, and
further research is needed into techniques for sys-
tem delay reduction. [5] Video see-through and
optical see-through head-mounted displays have
been the traditional output technologies for VR
applications for more than forty years. However,
they still suffer from several technological and
ergonomic drawbacks which prevent them from
being used effectively in all applications. [1] The
use of different types of AR displays will allow a
wider opportunity for developers to create new
applications and improve on the existing hard-
ware. The large amount of academic and com-
mercial research on AR entertainment will begin
the forefront for new devices in the near future.
[19]

6 CONCLUSION

This paper analyzed the history, technical as-
pects, and future trends of Immersive first-
person Virtual Reality Head-Mounted Displays.
From the form factor starting at the necessity of
being tethered from the ceiling, to finding a small
but powerful microdisplay for a fully immersive
environment. We also examined which sensors
are involved with computing the orientation of
the user at any point in time. These sensors
gathered data which was computed and given to
an application in which the Kalman Filter Algo-
rithm was applied to remove any unwanted er-
rors in the data in order to reciprocate an accu-
rate an efficient constant orientation of the head
of the user. To lightly state it, an exciting era
of new technologies is about to emerge driven by
mobile wearable displays as it applies to our daily
dives in the same way portable phones are glued
to the ears of billions of people, as well as to
high tech applications such as medical, deploy-
able military systems, and distributed training
and education. [12]
7 REFLECTION

My research on immersive first-person virtual reality did not start as a blank slate. There have been several courses and experiences that I have had in the past 4 years which have helped in the development of this project. The following courses have attributed the most to the project: Computer Architecture, Discrete Computational Structures, Computer Graphics, and Data Structures. The field of virtual reality is heavily dependent on the improvement of its hardware. Computer architecture allowed greater insight on what hardware is used within head-mounted displays. Hardware sensors are one of the key technical components with strong emphasis on math based sensor algorithms. Discrete Computational Structures, although not the type of math that is involved with sensor algorithms in head-mounted displays, gave me a better understanding on how computers and math are so intertwined. The digital displays within the headsets are key to the immersive experience. Computer Graphics shed light on how virtual reality applications are limited by today's computer graphic technologies. Finally, Data Structures gave me a wider view of how the applications are programmed and efficiently utilized.

References


