AUGMENTED AND VIRTUAL REALITY

Martin Levek

Faculty of Mechanical Engineering, Brno University of Technology
Institute of Automation and Computer Science
Technicka 2896/2, Brno 616 69, Czech Republic
200487@vutbr.cz

Abstract: *In this paper we summarize our knowledge about virtual reality (VR) and augmented reality (AR). Specifically it's applications in industry, healthcare, entertainment, retail, training, education and others. How it can be controlled through a human-machine interface. Problems and limitations that are being worked on - like how to calculate collisions of virtual objects, nausea caused by latency and so on.*

Keywords: *Augmented reality, virtual reality, human-machine interface, industry 4.0, computer-assisted surgery, simulation.*

1 Introduction

The idea of virtual or augmented reality is hardly new. It has been here for decades. But in recent years our technology progressed far enough to allow the average person to be able to enjoy it from the comfort their own home. While it has mainly become popular for entertainment purposes that does not mean it is just for fun. It is also a pillar of Industry 4.0 alongside big data, cloud computing, 3D printing, the internet of things and others. Further uses can be found in healthcare, education, art and many others.

2 Applications

Both virtual and augmented reality are technologies with a wide ranges of application. We will go over some of them.

2.1 Industry

There are two most common applications for VR and AR in industry - first is research and development where it is used to review projects with a view that no simple computer screen can provide, to verify that assembly and disassembly are possible and of course to create new designs in an immersive environment. Second is as a marketing tool to boost sales. [1]

2.2 Healthcare

Computer-assisted surgery is a fast growing field. Virtual reality can be used to perform surgeries remotely, for simulating upcoming operations or for training. Augmented reality can help the doctor visualize the inside of the body. Tailor-made 3D model of the patient's anatomy is created from scans in specialized software, then it is given physical properties. Digital models of living tissues which have to update in real time require a lot of computational power. This is not an issue if used for planning - which can be essential in improving the patients odds of survival by reducing the time required for the surgery or minimizing intrusiveness of the surgery. [1]

2.3 Entertainment

For many the most obvious and interesting application. Even if you're not one of them, simply dismissing this aspect of AR and VR would be short-sighted. If nothing else, it generates revenue and investment, driving the technology forward. Game developers have been trying to make their products more and more realistic, and it doesn't get much more realistic then VR. Real movements transferred into the virtual world, turning one's head to change the point of view. There can also be health benefits compared to the traditional way of playing video games, for example by using an omni-directional treadmill for movement. AR can also be used for entertainment as the popular Pokemon-collecting phone application has shown - motivating many young people to go outside and catch them all.

2.4 Retail

Technology - specifically the internet - is quickly changing our lives and one obvious example is retail. Shopping from the comfort of our homes would have been a strange notion just a few decades ago, but is completely normal today. However, the innovation doesn't end there. V-commerce, short for virtual commerce, might change the retail landscape. Shopping in virtual reality is more immersive than through a computer screen and more convenient and personalized than physically going to a store. V-commerce is currently in the early adoption phase, but already paying off to some - for example clothing or jewelry retailers who developed AR apps for their customer's phones. After choosing whatever they're interested in, they simply use their phone as a mirror in which they can see themselves wearing said item. Similar technology is used by IKEA. Their app lets people see how their new furniture would look, and fit, in their home. [3]

2.5 Training and education

Using VR instead of hands-on training has several benefits. Removing risk and saving resources are two of them. Training pilots for example - saves fuel, removes the need for maintenance and the risk of an accident caused by an inexperienced pilot. We can also create scenarios which would be rare in the real world and thus impractical or impossible to be used for training, like natural disasters or big accidents. The learning environment can be precisely controlled. Virtual humans can be used as substitutes for real ones, further reducing cost. When training with actual humans, there is no need for them to be in the same location. This way of training also makes it easier to collect data, which can be used to improve the learning experience or to tailor the learning process to a persons individual needs. We can also stray from trying to be as realistic as possible and use some interesting functionalities, such as slowing down or speeding up time, making objects transparent, changing colors to see something we normally wouldn't be able to and so on. [1]

2.6 Remote assistance

Your car broke down and you have to call a professional and wait on the side of the road for who knows how long? Maybe you don't. Simply put on your AR glasses, have said professional (or an artificial intelligence) connect to your device and guide you through the process of fixing your car remotely - saving time and resources. Many similar scenarios can be found in industry or at home. Not only will you save money and time, but you might also learn something new. This could also be life-saving in the case of medical emergencies, especially in hard to reach locations.

2.7 Art

VR and AR offer unique ways to create art. Painting in three dimensions rather than on a flat surface is one of them. Also, what better place to create a virtual environment than in virtual reality. [2]

2.8 Others

Some other uses include urban planning and architecture, for giving the designers a better idea of how the finished project is going to look like. Virtual tours through museums, cities or caves. Virtual meetings, classrooms or concerts. Augmented reality in a car, highlighting important information such as the speed limit or showing navigation on the windshield to remove the need to take one's eyes off the road. [2]

3 Human-machine interface

An important part of technology like AR or VR is how to give inputs to the machine and receive outputs back. Some ways are not as straight-forward as a keyboard and a mouse and have their issues, which will hopefully be solved in the future. There are many ways to interface with you device: screen, sound (music, tones, voice), gestures, buttons, trackball, joystick, haptics, tracking of location, breath rate, eye movement, pupil dilation, heart rate, brain waves, even direct brain-machine interface which is being developed by companies like Neuralink. Making the interface ergonomical is crucial. If it feels uncomfortable it can be immersion-breaking and undesirable to use, especially for long periods of time.

A traditional solution like a mouse and a keyboard or a joystick is time-tested, works well, is familiar to users - overall a good choice, but definitely not the only one. Optical, magnetic or ultrasonic tracking are also options. Not having anything physical to touch when controlling your device may not feel very natural, so the virtual keyboard can be projected onto a physical surface. Body tracking is an option that involves categories like: hand tracking, hand pose recognition and eye tracking.

With hand tracking, sensors on gloves, computer-vision or other ways of tracking the movement and pose of the hands are used. Inputs from these sensors are mapped to a digital skeleton, and then interpreted by the device. This allows for very natural control over the environment, like picking up objects, pushing buttons and so on.

Hand pose recognition is easily confused with hand tracking, but it is not the same. Specific hand poses can be recognized by the computer, similarly to sign language. Each of these poses is programmed to trigger a certain event. This is obviously not as natural and immersive as hand tracking and requires user tutorials, it can be especially confusing if different programs use different poses for similar commands. The upside is that it requires less individual calibration and isn't as processor intensive.

Eye tracking is not ideal as the main input, as it is rather tiring. But it is a great way to quickly determine what the user is interested in (often faster than the user is aware themselves).

Visual interface: virtual environment can be projected into each eye differently (stereoscopic), making it more immersive and giving the user better depth perception. This can be done on one screen with the two images separated by polarization or temporal frequency, or with two screens. This sadly puts more strain on the eyes, and thus causes visual fatigue, which should be avoided or mitigated, especially if the device is being used for prolonged periods of time. Fusing the two images in the brain of the user may not be possible depending on the distance of the screen from the eyes and some other factors. To avoid this issue, whatever the user is looking at, must be in the "Panum's fusional area". If there is only one image for both eyes it is called monoscopic.

The delay between an action in the real world and its visual effect on the virtual world is called latency, it arises because of processing times, refresh frequency and data transport (which is even higher if the processing happens elsewhere and the image is streamed). This delay can cause nausea, in this case sometimes referred to as "cybersickness". Even though a classic computer screen is not what most people imagine when they hear virtual reality, a curved screen surrounding the user or an array of normal flat screens is also an option.

Important aspect of a visual interface is it's field of vision. The before mentioned screen array surrounding the user would not reduce the natural field of vision at all, while in the case of augmented reality glasses Mirosoft HoloLens, it's users reported their field of vision to be about 34°. If the user can see the virtual and the real world at the same time, it will likely cause cybersickness, so if the field of vision doesn't cover everything the user can see, it is most likely a good idea to block out the rest of their vision. The discomfort goes up with the increasing size of the field of vision - this is thought to be caused by peripheral vision and it's sensitivity to movement.

Screen quality - refresh frequency, image resolution, brightness or range of color all play a role in the users perception and potential discomfort. Resolution that can be perceived by the human eye is higher than that on most (if not all) current headsets. This could cause issues, especially when focusing on small details, the user will instead see the individual pixels, breaking the immersion and causing visual fatigue. Low refresh frequency is obviously bad, but high frequency negatively influences latency - both should be avoided.

Through sound: those of us who tried controlling a device through speech recognition may know how frustrating it can sometimes be to try to get it to understand, only to give up and and do so manually, may not be too thrilled about this prospect. There is also the issue of misunderstandings even if the computer understood the words themselves, they may have different meanings in different scenarios, and let's not even get into things like metaphors or synonyms. Significant breakthroughs in artificial intelligence may be required to make this way of communicating with computers seamless and natural. But the technology is getting better and on a device like AR glasses may be a decent choice, as it is a hands-free input, which is a big plus, and can therefore be used in combination with other inputs. Listening on the other hand is something without which VR and AR would simply feel incomplete.

Haptic: a virtual world can feel a lot more real if you could touch it, and you can, in a way. With haptic gloves, that stop your fingers from moving any closer together, after you grab a virtual object, the VR/AR can become incredibly immersive. Even something like temperature can be felt through a haptic interface.

Brain-computer interface (BCI) is a system that translates brain activity into inputs that the computer can understand. This is usually done by electroencephalography, where sensors are placed (non-intrusively) on the users head, but in the case of something like Neuralink (which is still in development and not yet commercially available), the electrodes are implanted straight into the brain, this method is called electrocorticography. Another method is magnetoencephalography (we seem to have found a pattern of unpronounceable terminology), which measures magnetic currents. By measuring oxygen concentrations in the brain, we can find what areas are more active. The last two methods unfortunately require magnetic resonance imaging, which is unlikely to be affordable for every household any time soon.

BCIs can't read people's minds, only recognize blurry concepts - if you for example wanted to make a fist with your left hand, the BCI would know that something is going on with your left hand, but this is about as specific as the current technology can manage to recognize. So it may not be very well suited for complex tasks, but it is a hands-free input, and that is always a big upside. The method which requires surgery and uses electrodes implanted straight into the cortex is much more accurate, as it measures activity of individual neurons, but still no mind reading. [2, 1, 4]

4 Challenges

Creating a realistic virtual environment that users can interact with is hard enough for normal computers. Doing so for VR and AR adds another layer of complexity. Augmented reality is an overlay on the real world, so we have to not only keep track of virtual objects, but real ones as well, all while synchronizing our reality with other users. All of these calculations on a device that can be carried may not be very realistic if we want to maintain a decent frame rate and battery life. Cloud computing can help, but it's not a silver bullet, it introduces issues like delay and requires signal. Delay and frame rate are especially important because the human body is not used to having a delay between its actions and affecting the world. If it is introduced, it can cause nausea. One way to avoid this issue in VR is to use "teleportation" for movement, also solving the issue of not having enough space to move normally, but at the price of being less realistic.

Physical phenomena that can be simulated include: Solid mechanics, with rigid bodies, poly-articulated solids or deformable solids. Fluids and particles, which greatly increase computational requirements. Topological changes such as fractures. Or changes of materials properties - like viscosity.

A big challenge is detection of collisions. If we were to check for collisions for every object against all other objects, the complexity of the calculation would grow quadratically with the number of objects. Needles to say, this is something to be avoided. So we employ optimization algorithms to reduce the complexity of this calculation.

There are two types of detection methods: discrete and continuous. In the discrete methods the algorithm is used in fixed time intervals and ignores what happens in between. The continuous methods on the other hand look for the exact moment of collision. This method is more suited for situations where it is necessary to avoid all collisions. The computing time is naturally higher for these methods.

Objects can be either convex (any two points on the object can be connected by a line which is entirely contained within the object) or non-convex. Collisions of convex objects can be detected by simpler, and thus faster, algorithms. When it comes to non-convex objects, we have two options: either to divide the non-convex object into multiple smaller convex objects, or use a more complex algorithm.

Problems can be divided into two types: two-body problems and N-body problems. There is a single moving object in a two-body problem which makes calculations easier. In an N-body problem all objects can move, dramatically increasing the complexity.

When calculating collision detection, geometric data is gathered, in first phase, object pairs that cannot enter into collisions are filtered out. Next is narrow phase where more precise calculations take place to detect potential collisions. In the exact phase very precise calculation of the interpenetration of objects tells us if and when the collision happened. A physical response based on this information is determined. Collision detection can fortunately be solved with parallel computations, making the process considerably faster. [1]

5 Conclusion

Virtual and augmented reality are technologies with a great potential and many uses some of which we probably didn't even think about yet, but they will become obvious in the hindsight. We have talked about the issues and challenges of these technologies, some of it's many applications, about ways to communicate with these devices through body tracking, visual interfaces - and how this can cause discomfort, fatigue or nausea. AR and VR have promising futures, and even though some of it might be just internet hype, we will be eagerly waiting for exciting new developments in this field of research.

References

[1] Bruno Arnaldi, P. G., and Moreau, G. *Virtual Reality and Augmented Reality: Myths and Realities.* Wiley-ISTE, 2018.

[2] Erin Pangilinan, S. L., and Mohan, V. *Creating Augmented and Virtual Realities: Theory and Practice for Next-Generation Spatial Computing.* O'Reilly, 2019.

[3] tom Dieck, M. C. *Augmented Reality and Virtual Reality: The Power of AR and VR for Bussiness.* Springer, 2018.

[4] Whyte, J. *Virtual Reality and the Built Environment.* Architectual Press, 2017.