THE IMPORTANCE OF EYE TRACKING IN AUGMENTED REALITY (AR) AND VIRTUAL REALITY (VR)

EVALUATING THE UNDERLYING BENEFITS OF EYE TRACKING IN AR AND VR USE CASES

SUMMARY

The VR and AR industries are still in their infancy; however, changes are happening within the industry at a record pace. Users and developers are both learning lessons about the immersive technology needed for the industry to grow. Certain features will become necessary for day-to-day use as the industry grows and matures into a mainstream market. Much like capacitive touch transformed the smartphone industry, technologies like eye tracking have the potential to do similar things in the VR and AR space.

Device awareness–a device's understanding of its position and environment in context with the user and content−is improving which in turn improves user experience by addressing users' needs sooner. Features such as touch display, GPS and barometer for location, gyroscopes, accelerometers for movement, fingerprint sensors for faster security, and voice for authentication and control have all improved device awareness. Moreover, eye tracking takes device awareness to a whole new level because it enables devices to understand where users spend their most valuable resource, their attention, at any given point in time. Device awareness brings a vast new source of data about the user's attention which can feed into the new attention economy where many different things are vying for a user's attention. Moor Insights & Strategy believes this is a new frontier in understanding user behavior and serving users exactly what they want to see and when they want to see it.

In this paper, we will cover the importance of eye tracking technology and how it enables better AR and VR devices via three key capabilities–foveated rendering, interpupillary distance (IPD), and user identification. We will also discuss how eye tracking has a profound impact on user experience and important metrics to consider when seeking an eye tracking solutions partner to develop a headset or an application for AR and VR.

EYE TRACKING ADOPTION

A strong trend towards both “world facing” and “human facing” sensors exist today as part of the urge to drive “device awareness”. The world facing sensors are outward facing sensors in cars like light detection and ranging (LIDAR), room tracking in AR and VR, and even smartphone cameras on the back of devices. Human facing sensors include Apple’s near infrared (NIR)/structured light sensors in the Apple iPhone X used for FaceID and Animojis, Microsoft’s Windows Hello for Face ID, Samsung’s Galaxy S9 retina scanner, and newer automotive solutions for driver monitoring.

Eye tracking is one of the signals that can be delivered by human facing sensors alongside Face ID, iris scanning, head pose, gestures, facial features, etc.

Eye tracking, along with user identification, may have the broadest and most profound impact of all human facing signals given that eye tracking measures human attention. The list of use cases could be long and cover all types of devices. Many commercial applications today help improve or test user experience including website design or packaging research. Security applications exist as well. PC OEMs like Acer, Alienware, and MSI recently adopted eye tracking in notebooks and monitors to enable new user interfaces and experiences that were not possible in the past. For example, eye tracking is used by eSports leagues to track a player’s gaze and improve the audience’s overall watching experience.

Some specialized advertising applications with user consent track users’ eye movements as they look at a webpage. Data can then show advertisers exactly how many users have seen an ad, noticed their brands, and consumed key marketing messages. There are also applications where test groups are equipped with wearable eye tracking devices when they walk through a store. Retailers then get useful information of what users saw and did not see and can optimize store layout accordingly. Advertisers and retailers find the analytics and insights gained from these eye tracking solutions invaluable in maximizing returns on their investments.

Apple, Facebook, Google, and Microsoft are all aggressively pursuing headsets as well as platforms for both VR and AR. We believe Apple’s ARKit and Google’s ARCore are laying the groundwork for a much larger AR and VR ecosystem for each platform. Microsoft’s commitment to immersive computing with Windows 10 is also driving the adoption of eye tracking technology. The introduction of the Windows Mixed Reality platform makes it more important for both businesses and consumers to adopt and adapt to VR and AR experiences. Microsoft recently announced the support of eye

tracking on their platform with the introduction of new eye tracking application programming interfaces (APIs). They also went a step further with the submission of a new industry standard for eye tracking in partnership with Intel, Tobii, and EyeTech DS.

Facebook acquired The Eye Tribe and Apple acquired SensoMotoric Instruments (SMI), which sparked further interest in eye tracking as an elemental key feature in this new industry. We believe these types of acquisitions can drive increased awareness of the technology's potential benefits. However, these acquisitions could also potentially delay the implementation of eye tracking as a feature as many former partners of the acquired companies no longer have a long-term solution for eye tracking. We believe this consolidation creates some uncertainty in the market, even though other independent players remain.

The smartphone industry is following Apple's example and designing in NIR/structured light sensors in large numbers of high end phones. These sensors are used today for Face ID and Animojis, but we believe with minor adjustments they could be used successfully in eye tracking technologies as well.

The automotive industry is considering using eye tracking to monitor driver condition. In fact, European regulators are considering it as a requirement for European New Car Assessment Program (NCAP) 2025 implemented in 2020[1]. There are also a wide array of specialty markets for eye tracking that already exist today like medical, robotic surgery, and prescription glasses.

Tobii, who we believe is one of the industry leaders in the eye tracking market, has decades of experience in eye tracking and enables all types of applications across multiple industries. Tobii's position as both a technology developer and integrator allow it to bring eye tracking to multiple industries and to expand into new ones like the AR and VR industry.

We will now dive into foveated rendering, eye position including IPD and user identification−three key areas where eye tracking can lead to better devices for AR and VR that will help move the industry forward.

FOVEATED RENDERING

Foveated rendering is a rendering process designed only to show a user a portion of what they are looking at in full detail. This technique acts on the scientifically proven fact

[1] Euro NCAP 2025 Roadmap, September 2017, https://bit.ly/2LEWoAB

that our eyes have a narrow field of view at full resolution and blurring occurs outside of the fovea, also known as the center of the retina. Rendering an entire screen in full resolution generates waste in the most scarce and expensive components of any computing system because our eyes do not see everything at full resolution. Graphics shown with foveated rendering better matches our true way of seeing objects and delivers several benefits. The three benefits listed below are interlinked and device manufacturers can choose where on the tradeoff curve they want to be.

- Improved image quality: Foveated rendering can enable 4k displays on the current generation graphics processing units (GPUs) or graphical settings can be improved without degradation in frames per second (FPS) or performance.

- Lower cost: The same games or applications can run on lower cost hardware without performance degradation, which means end users can play the same games on lower cost devices. For game designers or VR/AR manufacturers, this increases their addressable market since more computers can run their game and/or headset.

- Increased Frame Rate per Second (FPS): Games can run at a higher frame rate using the same graphical settings.

There are two types of foveated rendering: dynamic foveated rendering and static foveated rendering. Dynamic foveated rendering follows the user's gaze using eye tracking and renders a sharp image wherever the user's retinas are looking rather than at any fixed location. This type of rendering provides a better user experience and image quality for the user over static foveated rendering.

Static foveated rendering applies a fixed area of foveation where the highest resolution is situated at the center of the viewer's device regardless of the user's gaze. It usually follows the user's head movements, but image quality is drastically reduced if the user looks away from the center of the field of view.

Foveated rendering allows for the implementation of higher resolution displays in VR and AR headsets, but does not require the device to render the full resolution. Tobii's internal studies show that the number of pixels rendered can be reduced by a range of approximately 50 to 70 percent without the end user noticing a reduction. This saves considerable GPU resources which can be translated into one or several of the benefits listed above.

Many different hardware and software pieces integrate tightly to enable foveated rendering. Specifically, the latency and synchronization of the full image rendering chain is key. The eye tracking algorithms must be highly optimized and run on suitable processing hardware. The eye tracking system relays the gaze point information to the game engine, which needs to tell the GPU how to render the graphics correctly. This process must happen in milliseconds or the user wearing the headset will notice a lag between where they are looking and what is being rendered correctly. This tight integration of all components drives user experience and is an absolute requirement because of the sensitivity to lag.

Put simply, foveated rendering enables higher resolution display options that otherwise wouldn't have been available to headset makers. Early implementations of foveated rendering shows improved pixel shading performance of 2x-3x[2], from which all the above benefits derive.

EYE POSITION

When it comes to AR and VR headsets, the graphics quality is among the most important parts of the entire device. The device optics determine how a user sees the display inside of the VR or AR headset, which can determine how a user perceives overall quality. However, the headset makes small adjustments to compensate and optimize for eye position to deliver peak image quality, as not everyone's eyes are the same.

There are two key aspects of eye position in VR and AR. One aspect is measuring the distance between the eyes. This measurement is the IPD, whichvaries significantly between users and is an important factor in maintaining optical sharpness and image quality. With eye tracking sensors, IPD can be measured and automatically adjusted or users can be guided to their optimal setting. Automatic IPD adjustment solves one of the barriers to entry for VR and AR−the inability to maintain accurate lens alignment−and will make the graphical experience better.

In addition to IPD, the actual position of the eye relative to the display is critical. This is both the distance between the eye and display and if the head-mounted display (HMD) is tilted relative to the eyes, which sometimes happens during intense usage like gameplay. Having real-time knowledge about the relative position of the eyes versus the

[2] NVIDIA Research, "Perceptually-Based Foveated Virtual Reality", July 2016, https://bit.ly/2si0Q06

headset is important as image quality can significantly degrade if a user's head moves within the headset. The user may not even be aware of these slight movements.

With eye tracking, an application can guide a user to adjust the headset if it's out of position. The measurement of the eye to display distance is especially important for AR displays because they cannot render 3D objects correctly if the eye position is not near optimal. The headset needs to know where the eyes are to compensate and to adjust for the difference to optimize image quality and ultimately user experience.

Headset manufacturers can ensure users have the clearest image possible and best possible graphical experience, using eye tracking for both IPD and eye position relative to the lenses and display. This is particularly important and may be a clear differentiator as graphical experience is a key selling point for VR. Image quality allows users to lock on to their immersion for longer and experience greater enjoyment. Eye tracking enables better image quality, which can also translate to less eye strain in VR or AR as the user's eyes spend less effort trying to focus.

USER IDENTIFICATION

Apple, Microsoft, Samsung, and others have implemented different forms of biometric authentication to help increase ease of use and security of their devices. One increasingly popular method of identifying users is the use of iris scanning. This process occurs where an infrared light illuminates the user's eyes and an infrared camera reads its unique iris pattern. This method is difficult to spoof and may be just as or more secure than fingerprint biometric identification as it allows for secure user identification without them having to touch anything. Iris identification can be done with the same sensors as eye tracking so no additional hardware is needed.

VR and AR headsets can look inward toward the user with iris scanning and identify them without the user having to do anything other than put the headset on their head. This kind of user experience resolves a key pain point for AR and VR because it allows for multiple users to share the same headset. Users can also swap identities on the same device in a classroom or enterprise environment, for example. All users immediately get their unique profile and settings. The alternative is to have the user login with a virtual keyboard which takes time and is prone to mistakes.

The ability to quickly and securely identify a user opens other opportunities for VR and AR headsets. Users can have different profiles and automatic user configurations, which makes the headsets more user friendly, but also more enterprise-friendly. The

technology enables companies to securely provide content and documents to users in a unique way that is literally for “your eyes only”. Iris identification via eye tracking also allows a user to prove their authenticity as an original author of content if they create it in AR or VR.

EXPERIENCE IS PARAMOUNT

Ultimately, the most important purpose of eye tracking is to improve the user experience. We believe there are a few experiential capabilities that are gained using eye tracking, which help to improve the overall user experience beyond a headset without eye tracking. We also believe having these experiential improvements will help HMD manufacturers and application developers set themselves apart from the rest of the competition as more headsets enter and grow the market.

FIGURE 1: HAND-EYE COORDINATION

Source: Tobii

Hand-eye coordination can be improved with eye tracking by coordinating the gaze (aim) with the action (throwing) to more accurately represent how humans interact with the real world. Without eye tracking, developers need to determine if the user hits an

object only based on the information of the movement of the hand controller, for example. When eye tracking is combined with gesture control, developers get two pieces of information: (i) how the hand controller moves and (ii) where the person was looking at the same time. With these two pieces of information, more accurate predictions and game experiences can be created.

FIGURE 2: EYE CONTACT

Source: Tobii

Second, the small, but critical social interaction of eye contact is another experiential improvement. The user can gain a more life-like experience in VR with the ability to see another user and where that user is looking. Without eye tracking, characters in a multiplayer setting usually stare straight ahead since the game designer has no information about what the person is looking at. Game designers can guess where a person is looking, but guessing incorrectly results in an unnatural experience.

With eye tracking, multiplayer characters' eyes will replicate the eye movements of the users. NPCs (non-playing characters), which are traditionally powered by artificial intelligence (AI), can become more life-like with eye tracking and respond accordingly when another player looks at them, stares at them, or ignores them. Without eye

tracking, game designers must guess what a user is focusing on and NPC responses are guesswork and often break immersion.

FIGURE 3: REDUCTION IN INTERFACE STEPS

Source: Tobii

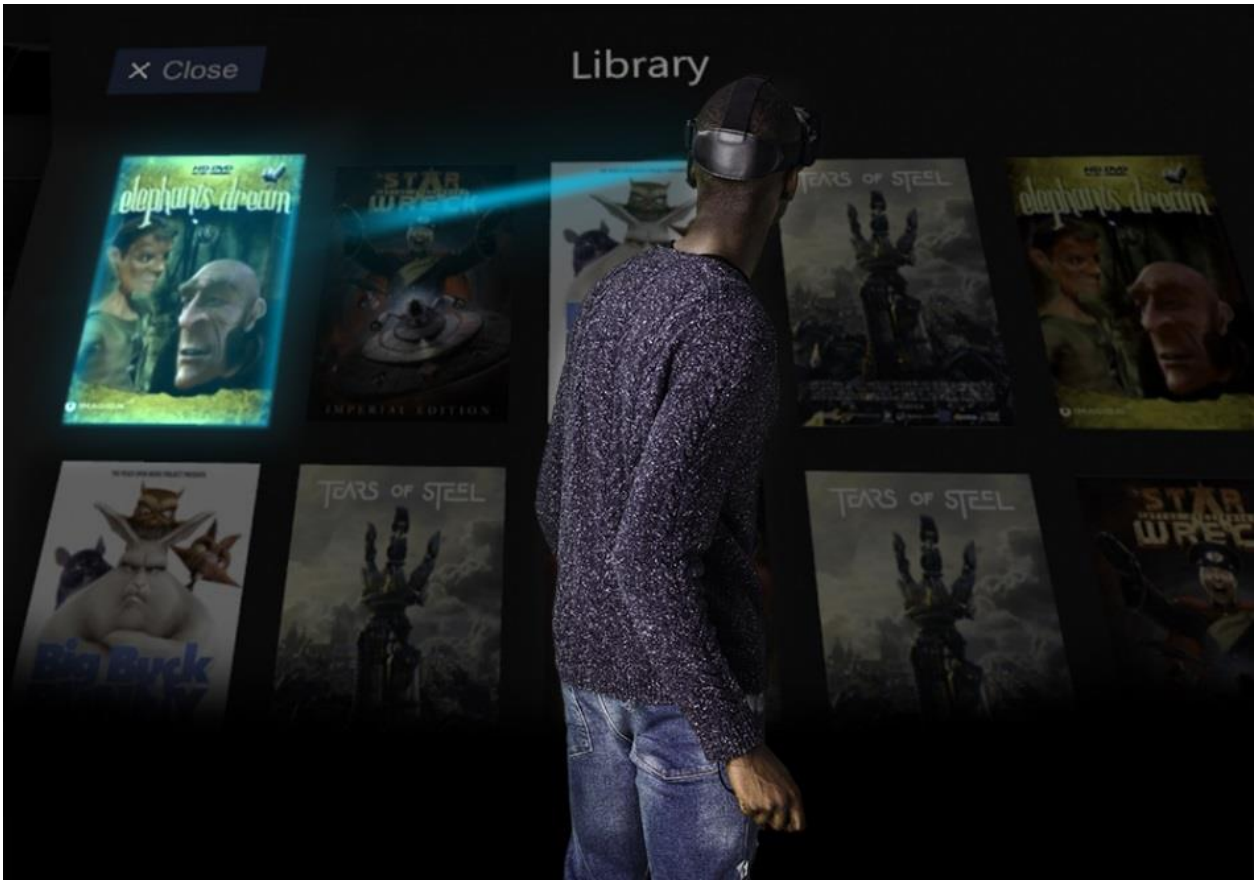
Eye tracking enables users to interact with objects or user interfaces with fewer steps. This change in user interaction is similar to how the touchscreen transformed smartphones and PCs from the traditional click and drag interfaces of old PCs. The difference amounts to quicker and more natural user interactions. Without eye tracking, interaction follows the below sequence:

1) Look at an object
2) Point at the object with your controller or forehead
3) Select by pressing a button.

With eye tracking, the interaction can be optimized to:

1) Look at an object
2) Select it by pressing a button

FIGURE 4: REMOVAL OF NON-INTUITIVE IMMERSION BREAKING INTERFACES



Source: Tobii

Lastly, eye tracking can help eliminate unnatural user interactions that break immersion and don't particularly exist in the real world such as using only head position to direct a pointer. In VR and AR, it has become common practice to point to objects with your forehead to select them. However, few examples exist in real life where humans behave that way. For instance, imagine you are the quarterback for an American football team and want to pass the ball to a person running full speed across the field. Aiming without eye tracking involves following the intended receiver with your forehead as they run, a cumbersome experience. This industry accepted immersion-breaking behavior in VR and AR experiences should be removed because it does not replicate reality.

IMPORTANT METRICS FOR AN EYE-TRACKING SOLUTION PARTNER

OEMs face difficulty and complexity when implementing eye tracking on multiple platforms with multiple users. To help with this challenge, certain metrics should be considered before picking an eye tracking partner. We believe the complexity of being able to support all types of users in all types of conditions is what sets different eye tracking solutions apart. This capability is measured in plots of accuracy over a total population and can be broken down by various aspects such as different gaze angles, for example.

One of the key questions one must ask is how much experience does the eye tracking partner have in optimizing hardware and algorithm configurations to work in all situations with all users?

Sourcing components for the hardware portion of the eye tracking solution and tuning them to operate in all scenarios and sub-optimal conditions is no small feat, requiring key knowledge and experience. Most eye tracking companies source their components from different trusted suppliers, while others partner with their component suppliers to create custom solutions that integrate the best for their clients' needs. The expertise in knowing when to integrate off-the-shelf components and when to use custom-built components is crucial to enable the best hardware solutions that work with the right algorithms.

Long term, eye tracking solutions need to be robust to address changes in the future. This means having a certain level of "future-proofing" as performance is expected to constantly evolve through software. We believe a key partner to address the constantly changing market would be strong in these six areas:

- Optical competence – the right knowledge in hardware and software components
- Algorithm competence – the experience and ability to constantly improve eye tracking algorithms and to address the different types of users that need support
- Multiple processing options – the ability to support both application-specific integrated circuit (ASIC) and integrated system on chip (SoC) solutions and to process at the lowest power possible with low latency
- UX competence – the experience and knowledge in how to build the best UX that utilizes eye tracking and meaningfully improves the UX

- Software applications – a proper set of APIs, software development kits (SDKs), and third-party apps that utilize eye tracking for solution longevity

- Deep technology research capabilities – the ability to discover better ways of delivering eye tracking and to make those discoveries available to partners

Aside from component selection and performance, sensor data must be processed with low-power and low-latency. ASICs are useful for this, but raw infrared camera sensor data in the headset can stream straight into a SoC's image signal processor. The ability to integrate both types of solutions with the right algorithms, software, game engines, and graphics drivers is important to ensure the platform works smoothly.

Few companies exist today that have experience in integrating eye tracking solutions into both commercial and consumer grade products as the AR and VR markets are relatively new. Both commercial and consumer segments expect to see rapid growth so being able to integrate solutions for both segments at scale are important. OEMs should weigh consumer and commercial experience and the ability to scale when selecting a partner.

It is also important to have scale at a global level and few suppliers do at present. As a result, many VR and AR headset OEMs partner with a relatively small number of larger suppliers. One of the reasons behind this is to have a global-scale partner ecosystem in place which means having software integration experience with different independent software vendors (ISVs) for different types of applications. For hardware solutions to be viable, applications are required to support the SDK of that eye tracking partner.

Long-term viability is another major factor to consider when choosing an eye tracking company. By our estimates, the AR and VR market still has years to go until they reach mainstream adoption and market size coupled with the usual organic churn of developing markets. Start-ups form and are acquired, others fail, and some succeed. OEMs need to consider whether a prospective partner has the resources, leadership, research investment, and experience to maintain a long-term partnership. Companies could face a risk in choosing the wrong partner and being suddenly left without a long-term solution.

There are many different players in the eye tracking space and a strong intellectual property (IP) position is crucial for any partner. Ensure you are tapped into the best IP; otherwise, you run the risk of infringing on another's IP if you create a solution that has already been patented. Having a partner with a strong IP portfolio can help minimize

risk, offer support in patent disputes, and provide opportunities to expand a solution's applications.

CALL TO ACTION

Companies in the AR and VR space should study the advantages of eye tracking technology and evaluate the experience and reliability of a partner before considering them for any technology integration. The benefits of having a partner with decades of experience in researching and developing eye tracking technology and the expertise in how to implement it both in hardware and software are enormous. We recommend developers and headset OEMs consider Tobii as their partner in implementing eye tracking solutions.

IMPORTANT INFORMATION ABOUT THIS PAPER

CONTRIBUTORS

Patrick Moorhead, Founder, President, & Principal Analyst at Moor Insights & Strategy
Anshel Sag, Associate Analyst at Moor Insights & Strategy

INQUIRIES

Contact us if you would like to discuss this report, and Moor Insights & Strategy will respond promptly.

CITATIONS

This paper can be cited by accredited press and analysts but must be cited in-context, displaying author's name, author's title, and "Moor Insights & Strategy". Non-press and non-analysts must receive prior written permission by Moor Insights & Strategy for any citations.

LICENSING

DISCLOSURES

This paper was commissioned by Tobii. Moor Insights & Strategy provides research, analysis, advising, and consulting to many high-tech companies mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

DISCLAIMER

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. Moor Insights & Strategy disclaims all warranties as to the accuracy, completeness, or adequacy of such information and shall have no liability for errors, omissions, or inadequacies in such information. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators and not as precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on what the future holds, they are subject to risks and uncertainties that could cause actual results to differ materially. You are cautioned not to place undue reliance on these forecasts and forward-looking statements, which reflect our opinions only as of the date of publication for this document. Please keep in mind that we are not obligating ourselves to revise or publicly release the results of any revision to these forecasts and forward-looking statements in light of new information or future events.