Lagging Human Factors – Latency in Helmet Mounted Displays and its Implications

Siddharth Shyamsundar

BAE Systems Operations Limited, United Kingdom

SUMMARY

Despite having significant human factors (HF) implications, latency in military aviation Helmet Mounted Displays (HMDs) is often seen as purely an engineering problem. This systematic review provides an overview of latency in HMDs, highlights the adverse impact it has on key HF and pilot performance metrics and outlines current latency thresholds and how they may not be relevant to the military aviation context. Drawing from extant literature, this paper provides recommendations for HF involvement, research and technology application in the development of low latency HMDs. It serves as a call to action, promoting significant HF involvement within a multidisciplinary, collaborative approach to exploring latency in the development of future military aviation HMDs.

KEYWORDS

Helmet Mounted Displays, Latency, Pilot Performance

Introduction

In military aviation, a key function of augmented reality (AR) devices like helmet mounted displays (HMDs) is to overlay virtual information over the real world. Virtual symbology and imagery showing flight parameters and sensor information are displayed on a see-through visor. The pilot can attend to this information while still looking into the outside world through the visor.

A key issue plaguing AR HMDs is registration, which is the correct alignment of virtual information with the real world. Any movement by the user or in the scene introduces issues with temporal and spatial alignment, which may prevent the seamless augmentation of virtual information with the real world (Itoh et al., 2021). This misalignment is magnified in outdoor conditions with uncontrolled ambient lighting and movement (Azuma, 1997; Bibb, 2019). In a study by Roberts et al. (2013) where a soldier used an HMD in an outdoor environment, misalignment was found to cause the following behaviours in a virtual object:

- Jitter: A high frequency fluctuating motion.
- Wander: A low frequency movement observed over minutes and hours.
- Lag: A temporal delay in the motion of a virtual object in response to real world motion.
- Bounce: A “jumping”, vertical movement of a virtual object when walking.
- Accuracy: An angular error between a virtual object and the real world element it overlays.

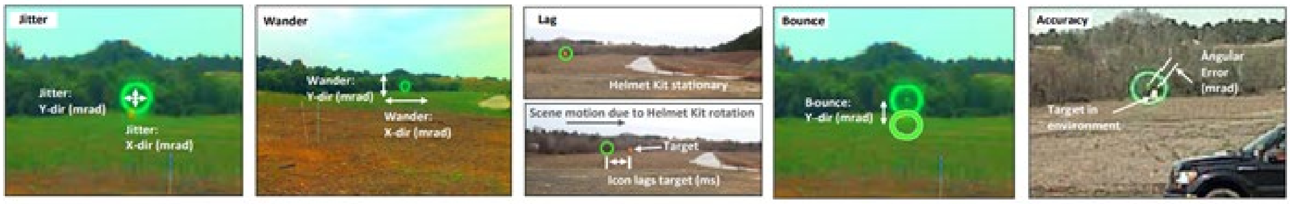

Figure 1: Behaviours exhibited by a virtual object as a result of misalignment (Roberts et al., 2013)

Latency is the misalignment caused due to the delay between an action (user or scene movement) and the system's response. It is conventionally measured in milliseconds (ms). With AR HMDs used in military aviation, latency refers to the time taken between the pilot moving his/her head and the updated symbology appearing on the visor. It can also occur due to the lag between the movement of features in the real world and the corresponding changes being displayed to the pilot (Stauffert et al., 2020). This can be observed when using a synthetic image while flying at low level in low light conditions. It is vital that the synthetic image updates rapidly with outside world and ground terrain information. Any temporal or spatial latency in the HMD while rendering this image may prevent the pilot from making the correct control and navigational inputs, posing a threat to pilot and aircraft safety and could result in potentially disastrous consequences.

Latency in Military Aviation HMDs and Human Factors

Latency in military AR HMDs poses a litany of human factors (HF) and performance challenges. Interestingly, one of the reasons for this is the way latency interacts with our sensory systems. Human perception of motion and spatial orientation is detected and processed by the vestibular and peripheral visual systems (Hosman, 1997; Meiry, 1965). These provide us with a sense of balance and stabilise our line of sight during head movement using the vestibulo-ocular reflex (VOR) (Fetter, 2007; Melzer, 2017). As latency causes visual cues to lag behind other perceptual cues of the vestibular system, it impedes the VOR and induces sensory conflict, resulting in fatigue, disorientation, physical discomfort, nausea, and simulator or motion sickness (Jerald, 2009).

These effects are more likely and more acute when using an HMD with a wide field of view (FOV) (Bailey et al., 2004; Buker et al., 2012; Ebenholtz, 1986; Jennings et al., 2004; Johnson, 2005; Melzer, 2017; Reason & Brand, 1975). Current fixed wing military HMDs have a maximum FOV of 40° and provide information to the foveal area of human vision (BAE Systems, 2023; Collins Aerospace, 2023). While this area processes information for higher cognitive tasks, it is very poor at sensing motion and orientation. As a result, these HMDs do not engage the vestibular and peripheral visual systems, and this in turn limits the impact increased latency may have on the pilot. However, the use of wide FOV HMDs will engage these systems, making latency more apparent and exacerbating its ill-effects on pilot performance. Given that wide FOV HMDs may become commonplace in the future, it is critical to explore latency's impact on pilots using wide FOV HMDs, along with setting up mitigation strategies and appropriate technical requirements.

Visual latency in HMDs has been found to have significant adverse impacts on task performance across a range of basic tracking tasks, complex skilled motor tasks and collaborative tasks (Ellis, 2009; Gunn et al., 2005; Sprague et al., 2006). Middendorf et al. (1991) reported larger altitude errors and higher workload when those performing simulated low level flying experienced visual lag. Latency can also result in pilot induced oscillations and impact tasks requiring high frequency, high precision inputs based on synthetic imagery from the HMD like night time air-to-air refuelling and low level flying. Based on a helicopter flying task, Jennings et al. (2004) further highlighted adverse effects of visual latency such as higher workload, poorer aircraft handling and increased positional errors. Preventing and mitigating these issues is key to safety and mission effectiveness.

Latency can negatively affect the user's sense of presence and immersion, both of which enable effective performance by providing a sensation of "being there" and experiencing the virtual content as continuous and one with the real world (Slater et al., 2009; Stanney et al., 1998; Steptoe, 2014). Latency tends to cause an invasive disruption of presence, regardless of how high the fidelity of the rendered virtual scene is (Meehan et al., 2003; Zimmons & Panter, 2003). The mismatch between head motion, environmental movement, the real world and the augmented image can also cause Oscillopsia, whereby the user may perceive the world to swim, oscillate and move. This can result in a reduced sense of presence, spatial disorientation and nausea (Allison et al., 2001).

Along with FOV, the perception and impact of latency is moderated by the presence of motion cueing. Bailey et al. (2004) found that participants were more sensitive to and perceptive of latency in conditions with motion cues. They reported that in low latency assessments, pilot performance was better with motion cues than without. Similarly, in high latency assessments, pilot performance was worse with motion cues than without. They concluded that motion cues enhance performance in low latency conditions, but more adversely impact performance in high latency scenarios. This has key implications for HMD development in military aviation. Here, motion cues from the outside world, either seen by the eye in the day, or through imagery at night, will always be visible to the pilot and may thus call for low latency requirements in the HMD to enable pilot performance.

The effects of latency also depend on the task involved, the amount of head movement, safety requirements and the amount of scene motion. For search and track tasks, where users may move their heads quickly, even small amounts of latency could cause impaired performance, discomfort, fatigue and nausea. Bailey et al. (2004) stated that for time-critical tasks and those that involve high levels of precision, low latency is vital for safe and effective performance. However, for tasks with more time and greater allowances, higher levels of latency may be permissible.

System and data complexity play a very important role. When an HMD displays imagery from multiple sensors, latency perceived by the user may be affected by the different latencies of each sensor. An HMD showing imagery from a night vision camera, primary flight symbology and weather data might be impacted by the four different latencies of the head tracker, night vision camera, the symbology system and the weather radar. These individual latencies may themselves be moderated by scene complexity and motion. Synchronising all these latencies in the system is vital, and this can be done using dynamic stream synchronisation. This is a system that continuously monitors latency values of each sensor or data stream and synchronises them such that they are displayed to the user at the same time (Azuma, 1997; Jerald, 2009; Kijima & Miyajima, 2016).

Clearly, these findings underline that latency tends to have far reaching and significant effects not just on the usability of the AR HMD, but critically, also on key HF and performance metrics and safety. Despite this, latency has often been seen as purely an engineering problem and research on latency requirements for AR HMDs appears to be very divisive and with limited consensus.

Latency Thresholds and Requirements

Several studies have tried to establish minimum latency requirements, thresholds above which latency in an AR system becomes perceptible. While these refer to AR head mounted systems in general, they play in important role in informing latency requirements for AR HMDs in military aviation. Blate et al. (2019) studied latency using an ultra-low delay system and called for AR displays to have latencies below 1ms. Jerald (2009) argued that latency should not exceed 3ms, while a number of studies have recommended that AR displays have a maximum latency of 5ms (Daqri, 2018; Itoh et al., 2021; Jerald & Whitton, 2009; Livingston & Zhuming, 2008).

Mania et al. (2004) found that users perceived latency at 15ms. The United States Department of Defence's (2012) Military Standard 1472G requires HMD latency to not exceed 16ms. Adelstein et al. (2003) recommended a threshold of 17ms, while Bailey et al. (2004) called for a latency of 20ms for demanding tasks in high resolution, wide FOV HMDs. There clearly appears to be a wide range across recommendations and minimum thresholds for AR HMDs.

A possible reason for such a broad spread across the recommendations is that individual latency detection thresholds vary widely, ranging from 3ms to over 100ms (Jerald, 2009; Jerald & Whitton, 2009). This, along with the fact that the human-visual perceptual moment (the smallest unit of time a human can perceive) lasts 60ms to 70ms supports recommendations of 50ms to 70ms (Albert et al., 2017; Coren et al., 1999; Efron, 1967; Jennings et al., 2004). However, latency thresholds should be lower than perceptual moments as one can still suffer from impaired performance and the other ill-effects of latency without having to perceive the latency itself (Jerald and Whitton, 2009).

Table 1: Overview of different latency thresholds recommended in literature

Latency Threshold Recommended (in ms)	Study
1	Blate et al. (2019)
3	Jerald (2009)
5	Daqri (2018); Itoh et al. (2021); Jerald & Whitton (2009); Livingston & Zhuming (2008)
15	Mania et al. (2004)
16	US Department of Defense Military Standard 1472G (2012)
17	Adelstein et al. (2003)
20	Bailey et al. (2004)
33	Rash et al. (2009)
50 – 70	Albert et al. (2017); Coren et al. (1999); Efron (1967); Jennings et al. (2004)

Several latency studies (especially those recommending 16 – 17ms) are limited by use of 60 Hertz (Hz) displays. These refresh 60 times a second, presenting a new frame every 16.7ms. Hence, with these displays, the lowest latency that can be assessed is 16.7ms, with subsequent test points being multiples of this (Mania et al., 2004). Such displays have been used in several studies, including those by Adelstein et al. (2003) and Ellis et al. (2004), and may restrict the thresholds they generate. As technology develops, using 90 Hz and 120 Hz displays can provide further granularity and a wider range of test points, with lower base latencies of 11.1ms and 8.3ms respectively (Itoh et al., 2021). Moreover, the use of software like Netdisturb can add greater rigour to the generation of latency requirements. Netdisturb artificially creates latency at different user-definable levels, allowing for a much wider range of testable latencies (ZTI Communications, 2022).

Today, AR HMD prototypes can be built to very stringent specifications in order to assess latency. Itoh et al. (2016) built an AR HMD with a latency of less than 1ms, customising a commercial off the shelf (COTS) headset. More recently, Ishihara et al. (2023) demonstrated an HMD with less than 4ms of latency. These satisfy the latency requirements of 5ms that several studies have recommended (Daqri, 2018; Itoh et al., 2021; Jerald & Whitton, 2009; Livingston & Zhuming, 2008). Today, techniques such as foveated rendering and predictive head tracking show promise in reducing latency. However, how images are transmitted and rendered, the display technologies used and form factor all play a vital role in the overall latency of an HMD (Itoh et al., 2021).

Currently, most evidence points towards a latency threshold of 5ms to 16ms. However, the limitations posed by the use of narrow FOV displays, lower refresh rates, the lack of motion cueing and the individual differences in latency perception all bring these recommendations into question. It is vital to take into account the wide FOV, high resolution, motion cued, time and safety critical environment of future military aviation. As such, this research can be seen as an initial stepping stone, with more focused research and experimentation vital to the development of future HMDs.

Recommendations

Future Research

- Latency research should include wide FOV, high resolution HMDs, where latency and its effects are more apparent and pervasive. Understanding latency's impact in such systems is vital to determining latency thresholds and developing future military aviation HMDs.
- Latency assessments should include virtual and real world motion cueing and assess their impact on HF performance metrics such as task completion, navigational performance workload, fatigue, spatial disorientation and simulator sickness.
- Leveraging modern ultra-low latency prototypes will help study the phenomenon and its effects in greater granularity and provide more detailed, robust and relevant results.
- Latency assessments should consider task complexity and context, including the high precision, high frequency, time and safety critical tasks that dominate military aviation.
- As latency perception differs across the population, assessing baseline latency perception in military aviators can help develop better HMD requirements and inform selection.

Future Technological Applications

- AR HMDs showing imagery may require more stringent thresholds as imagery may cause latency to be more apparent and impairing. This is exacerbated with see-through imagery, where temporal and spatial misalignment become very obvious.
- Industry should leverage the ultra-low latency prototypes developed by research institutions and work to "productise" these for use in critical operations.
- When multiple sensors display information on an HMD, dynamic stream synchronisation must be highlighted as a critical, core component in the requirements gathering phase.
- Technologies like foveated rendering, eye tracking and predictive head tracking must be exploited to reduce latency in AR HMDs.

The Role of Human Factors Specialists

- Engaging with pilots and operators through interviews and focus groups to better identify, understand and explore the HF problems arising as a result of latency in today's HMDs.
- Driving task analysis to identify elements of operational, tactical and domestic flying that are likely to be impacted by latency in HMDs and to work to mitigate the underlying risks.
- Designing appropriate human in the loop assessments to determine the impact of latency on key HF metrics like task performance, workload, response times and safety margins.
- Leveraging the assessment data to generate evidence based, contextual latency thresholds.

- Driving extensive studies to determine latency perception thresholds in military aviators to inform latency requirements, selection and human machine interface design.
- Upholding a duty of care to military aviators and establishing adequate mitigation strategies to ensure that the latency in AR HMDs does not adversely affect their health.

Conclusion

AR HMDs are a key cornerstone of competitive military advantage. With the pervasive effects of latency on usability, pilot performance and safety, there is a clear need for increased research, focusing on latency perception, better assessment methods, and improved thresholds.

Despite its implications on human performance, latency has often been viewed as an engineering problem. However, it is evident that HF and performance specialists should be intrinsically involved in the development, design and assessment of latency requirements for HMDs. The design of a latency-free HMD of the future requires a multi-disciplinary approach, driven by collaboration between academia and industry and consisting of engineers, researchers and HF specialists.

References

Adelstein, B. D., Lee, T. G., & Ellis, S. R. (2003). Head tracking latency in virtual environments: Psychophysics and a model. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *47*(20), 2083–2087. https://doi.org/10.1177/154193120304702001

Albert, R., Patney, A., Luebke, D., & Kim, J. (2017). Latency requirements for foveated rendering in virtual reality. ACM Transactions on Applied Perception, 14(4), 1–13.

Allison, R. S., Harris, L. R., Jenkin, M., Jasiobedzka, U., & Zacher, J. E. (2001). Tolerance of temporal delay in Virtual Environments. Proceedings IEEE Virtual Reality 2001.

Azuma, R. T. (1997). A survey of Augmented Reality. Presence: Teleoperators and Virtual Environments, 6(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

BAE Systems. (2023). Striker® II Digital Helmet-Mounted Display. https://www.baesystems.com/en-uk/product/striker-ii-digital-helmet-mounted-display

Bailey, R. E., Arthur III, J. J., & Williams, S. P. (2004). Latency requirements for head-worn display S/evs applications. Enhanced and Synthetic Vision 2004.

Bibb, C. (2019). Determining principles for the development of mixed reality systems for command and control applications. University of Birmingham.

Blate, A., Whitton, M., Singh, M., Welch, G., State, A., Whitted, T., & Fuchs, H. (2019). Implementation and Evaluation of a 50 kHz, 28µs Motion to-Pose Latency Head Tracking Instrument. IEEE Transactions on Visualization and Computer Graphics, 25(5), 1970–1980.

Buker, T. J., Vincenzi, D. A., & Deaton, J. E. (2012). The effect of apparent latency on simulator sickness while using a see-through helmet-mounted display. Human Factors: The Journal of the Human Factors and Ergonomics Society, 54(2), 235–249.

Collins Aerospace. (2023). F-35 Gen III helmet mounted display system. https://www.collinsaerospace.com/what-we-do/industries/military-and-defense/displays-and-controls/airborne/helmet-mounted-displays/f-35-gen-iii-helmet-mounted-display-system

Coren, S., Ward, L. M., & Enns, J. T. (1999). Sensation and perception. Harcourt Brace College Pub.

Daqri. (2018). Motion to photon latency in Mobile AR and VR. Medium. https://medium.com/@DAQRI/motion-to-photon-latency-in-mobile-ar-and-vr-99f82c480926

Department of Defense. (2012). MIL-STD-1472G. Department of Defence Design Criteria Standard. *Human Engineering, Department of Defence*

Ebenholtz, S. M. (1986). Properties of adaptive oculomotor control systems and perception. Acta Psychologica, 63(3), 233–246. https://doi.org/10.1016/0001-6918(86)90045-4

Efron, R. (1967). The duration of the present. Interdisciplinary Perspectives of Time, 138(4), 713–729.

Ellis, K. (2009). Eye tracking metrics for workload estimation in flight deck operations. University of Iowa. https://doi.org/10.17077/etd.a773626l

Fetter, M. (2007). Vestibulo-ocular reflex. Neuro-Ophthalmology, 35–51.

Gunn, C., Hutchins, M., & Adcock, M. (2005). Combating latency in haptic collaborative virtual environments. Presence: Teleoperators and Virtual Environments, 14(3), 313–328.

Hosman, R. (1997). Visual-vestibular interactions in the perception and control of aircraft motions by the pilot. Modeling and Simulation Technologies Conference.

Ishihara, A., Aga, H., Ishihara, Y., Ichikawa, H., Kaji, H., Kawasaki, K., Kobayashi, D., Kobayashi, T., Nishida, K., Hamasaki, T., Mori, H., & Morikubo, Y. (2023). Integrating both parallax and latency compensation into video see-through head-mounted display. IEEE Transactions on Visualization and Computer Graphics, 29(5), 2826–2836.

Itoh, Y., Langlotz, T., Sutton, J., & Plopski, A. (2021). Towards indistinguishable augmented reality. ACM Computing Surveys, 54(6), 1–36. https://doi.org/10.1145/3453157

Itoh, Y., Orlosky, J., Huber, M., Kiyokawa, K., & Klinker, G. (2016). Ost rift: Temporally consistent augmented reality with a consumer optical see-through head-mounted display. 2016 IEEE Virtual Reality (VR). https://doi.org/10.1109/vr.2016.7504717

Jennings, S., Reid, L. D., Craig, G., & Kruk, R. V. (2004). Time delays in visually coupled systems during flight test and Simulation. Journal of Aircraft, 41(6), 1327–1335.

Jerald, J. (2009) Scene-Motion- and Latency-Perception Thresholds for Head-Mounted Displays. University of North Carolina.

Jerald, J., & Whitton, M. (2009). Relating Scene-Motion Thresholds to Latency Thresholds for Head-Mounted Displays. Proceedings. IEEE Virtual Reality Conference, 211–218.

Johnson, D. M. (2005). Introduction to and review of Simulator Sickness Research. U.S. Army Research Institute-Rotary Wing Aviation Research Unit. https://doi.org/10.1037/e456932006-001

Kijima, R., & Miyajima, K. (2016). Measurement of head mounted display's latency in rotation and side effect caused by lag compensation by simultaneous observation — an example result using Oculus Rift DK2. 2016 IEEE Virtual Reality (VR).

Livingston, A., & Zhuming Ai. (2008). The effect of registration error on tracking distant augmented objects. 7th IEEE/ACM International Symposium on Mixed and Augmented Reality.

Mania, K., Adelstein, B. D., Ellis, S. R., & Hill, M. I. (2004). Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity. Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization.

Meehan, M., Razzaque, S., Whitton, M. C., & Brooks, F. P. (2003). Effect of latency on presence in stressful virtual environments. IEEE Virtual Reality, 2003. Proceedings.

Meiry, J. L. (1965). The Vestibular System and Human Dynamic Space Orientation. Massachusetts Institute of Technology. Department of Aeronautics and Astronautics.

Melzer, J. (2017). How much is enough? The human factors of field of view in head-mounted displays. SPIE Proceedings.

Middendorf, M., Fiorita, A., & Mcmillan, G. (1991). The effects of simulator transport delay on performance, workload, and control activity during low-level flight. In AIAA Flight Simulation Technologies Conference, New Orleans, LA (pp. 412-426).

Rash, C. E., Russo, M. B., Letowski, T. R., & Schmeisser, E. T. (2009). Helmet-mounted displays: Sensation, perception, and cognition issues. U.S. Army Aeromedical Research Laboratory.

Reason, J. T., & Brand, J. J. (1975). Motion sickness. Academic Press.

Roberts, D., Menozzi, A., Cook, J., Sherrill, T., Snarski, S., Russler, P., Clipp, B., Karl, R., Wenger, E., Bennett, M., Mauger, J., Church, W., Towles, H., MacCabe, S., Webb, J., Lupo, J., Frahm, J.-M., Dunn, E., Leslie, C., & Welch, G. (2013). Testing and evaluation of a wearable augmented reality system for Natural Outdoor Environments. SPIE Proceedings.

Slater, M., Lotto, B., Arnold, M. M., & Sanchez-Vives, M. V. (2009). How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicología, 40*(2), 193–210.

Sprague, D., Po, B., & Booth, K. (2006). The Importance of Accurate VR Head Registration on Skilled Motor Performance. Proceedings of Graphics Interface 2006, 131–137.

Stanney, K. M., Mourant, R. R., & Kennedy, R. S. (1998). Human factors issues in virtual environments: A review of the literature. Presence: Teleoperators and Virtual Environments, 7(4), 327–351.

Stauffert, J.-P., Niebling, F., & Latoschik, M. E. (2020). Latency and cybersickness: Impact, causes, and measures. A Review. Frontiers in Virtual Reality, 1.

Steptoe, W. (2014). What is presence in immersive augmented reality? http://willsteptoe.com/post/91843919913/what-is-presence-in-immersive-augmentedreality

Zimmons, P., & Panter, A. (2003). The influence of rendering quality on presence and task performance in a virtual environment. IEEE Virtual Reality, 2003. Proceedings.

ZTI Communications. (2022, November 26). NetDisturb. https://www.zti-communications.com/netdisturb