

VR Remote Magnified Viewing System Using an Ultrafast Pan-Tilt Camera^{*}

Daisuke Tahara^{1,*}, Yuta Abe¹, Taishi Iriyama¹, Takashi Komuro¹, Kohei Shimasaki² and Idaku Ishii²

¹Saitama University, Saitama, Japan

²Hiroshima University, Higashihiroshima, Japan

Abstract

We present a VR remote magnified viewing system that allows users to observe a wide-angle live video from a remote site while simultaneously inspecting a target region in high resolution. Our system realizes "Optical Foveation," a concept inspired by the human visual system, which provides high-acuity vision in the fovea (center of gaze) and a wide field of view in the periphery. The system combines a wide-field camera for contextual overview and an ultrafast pan-tilt camera with a galvano mirror for a magnified, gaze-contingent view. This mirror-driven mechanism achieves millisecond-level response, instantaneously aligning the magnified view with the user's head motion and ensuring a seamless transition from context to detail. To mitigate VR sickness, which is often exacerbated by the lag and visual-vestibular mismatch in magnified views, our system displays the telephoto image only on user command. This user-triggered approach minimizes exposure to high-magnification motion and enhances comfort. We describe the system's architecture and report on a prototype implementation, with experimental results confirming its responsive, comfortable, and effective operation.

Keywords

Remote magnified viewing system, Ultrafast pan-tilt camera, VR sickness, Optical foveation

1. Introduction

VR remote viewing systems offer an immersive sense of presence for applications like remote tourism, technical assistance, and site inspection [1, 10]. A critical challenge in these applications is the need to simultaneously perceive a wide contextual overview and inspect fine details of distant objects with high fidelity.

Conventional approaches to this problem have major drawbacks in VR contexts. Hardware-based solutions using mechanical Pan-Tilt-Zoom (PTZ) cameras [9, 13] suffer from significant mechanical latency. Their physical actuation is too slow to follow a user's rapid head movements, causing a critical gaze-to-image lag that is a primary contributor to VR sickness, as extensive research confirms [12, 8]. On the other hand, software-based alternatives, such as foveated rendering which computationally prioritizes gaze points [6], or super-resolution techniques [5, 11], introduce their own challenges. These methods can incur significant computational costs, potentially introducing new sources of latency, and may compromise information accuracy required for precise tasks due to their estimative nature.

To resolve this trade-off between latency and resolution, we propose a system that implements "Optical Foveation." Inspired by the human eye, which combines a high-resolution foveal view with a wide peripheral view, our system uses two separate cameras. A wide-field camera provides a stable, low-magnification peripheral image, while an ultrafast pan-tilt camera provides a high-resolution telephoto image of the user's gaze point. The key is a low-inertia galvano mirror that steers the telephoto camera's line of sight with millisecond-level response [4, 2]. This enables the magnified view to be redirected virtually instantaneously, minimizing the head-motion-to-display mismatch that is amplified at high

APMAR'25: The 17th Asia-Pacific Workshop on Mixed and Augmented Reality, Sep. 26-27, 2025, Busan, South Korea

^{*}We use the latest ceurart style.

^{*}Corresponding author.

✉ d.tahara.417@ms.saitama-u.ac.jp (D. Tahara); y.abe.796@ms.saitama-u.ac.jp (Y. Abe); iriyama@mail.saitama-u.ac.jp (T. Iriyama); komuro@mail.saitama-u.ac.jp (T. Komuro); shimasaki@hiroshima-u.ac.jp (K. Shimasaki); iishii@robotics.hiroshima-u.ac.jp (I. Ishii)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

magnifications. Furthermore, we empower the user to toggle the magnified view, limiting the duration of tight head-image coupling and reducing discomfort.

Contributions. We make the following contributions:

1. A VR remote viewing system implementing Optical Foveation, which combines a wide-field camera for context and an ultrafast pan-tilt camera for gaze-contingent magnified detail.
2. A low-latency, low-inertia control loop that maps HMD rotation to galvano mirror angles, achieving virtually instantaneous redirection of the magnified view.
3. A user-triggered display policy that mitigates VR sickness by minimizing exposure time to high-magnification imagery, thereby reducing visual-vestibular conflict.
4. A prototype implementation and a qualitative evaluation demonstrating the system’s responsiveness and usability for comfortable remote inspection tasks.

2. System architecture and implementation

This section presents the overall design, camera module, optical parameters, and the control/display pipeline. Figure 1 provides a high-level overview: a wide-field camera supplies the wide-angle image; an ultrafast pan-tilt camera with a galvano mirror supplies the magnified image aligned to the user’s current gaze direction. The user wears an HMD; head rotation from the HMD sensors is mapped to galvano mirror angles and sent via UDP to control the magnified view. The user presses a controller button to display or hide the magnified view. This policy preserves situational awareness in the wide-angle view and shortens the time under high magnification.

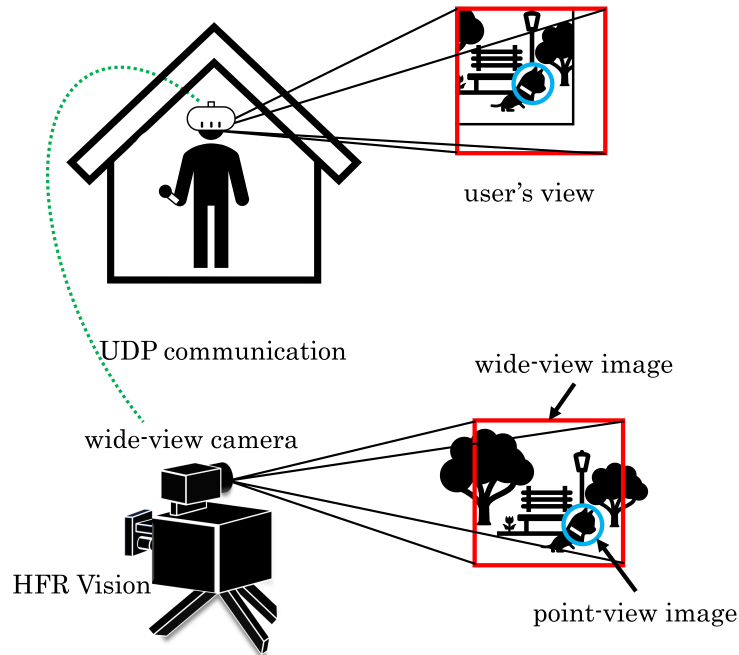


Figure 1: Overview of the VR remote magnified viewing system. A wide-field camera provides context; an ultrafast pan-tilt camera with a galvano mirror provides gaze-aligned magnification.

2.1. Camera module

Figure 2 shows the module, consisting of a 30 fps wide-field camera for target detection and a 120 fps ultrafast, mirror-driven pan-tilt camera for gazing. The galvano mirror enables smooth, low-inertia



Figure 2: Camera module with an ultrafast pan-tilt camera and a wide-field camera. The galvano mirror steers the magnified camera's line of sight at high speed.

redirection of the magnified camera's line of sight, avoiding bulk camera motion.

2.2. Optics and fields of view

The wide-field camera uses a 3.5 mm lens (horizontal FOV 70.7°, vertical FOV 56.6°); the ultrafast pan-tilt camera uses a 25 mm lens (horizontal FOV 8.5°, vertical FOV 11.4°). At a distance of 5 m, a 1440×1080 wide-angle frame covers 5.6×4.2 m (≈ 3.9 mm/px), while a 480×640 magnified frame covers 0.6×0.8 m (≈ 1.25 mm/px). Thus, the magnified view provides about $3.1\times$ finer linear sampling ($\approx 9.7\times$ per-area pixel density) than the wide-angle view.

2.3. Gaze redirection

Figure 3 illustrates the mapping from head motion to galvano mirror actuation. Head pitch/yaw from the HMD is converted to galvano mirror angles and transmitted over UDP, so the magnified view follows the user's gaze when enabled. Within VR, the user searches for a region of interest using the wide-angle view; once oriented to the target, pressing the controller displays the corresponding magnified view. Because the user chooses the timing, magnified display time is minimized, reducing overload and the likelihood of VR sickness.

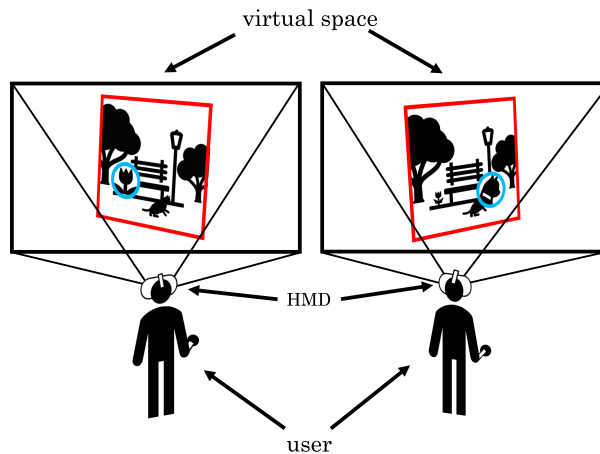


Figure 3: Switching gaze direction by mapping HMD head rotation to galvano mirror angles. The magnified view follows the user's gaze when enabled.

2.4. Display policy and composition

Figure 4 shows the view composition. The HMD shows the wide-angle view from the wide-field camera; the magnified view from the ultrafast pan-tilt camera is layered on top when requested. Even when only the wide-angle image is displayed, a semi-transparent indicator shows the prospective magnified region to aid targeting and shorten acquisition time.

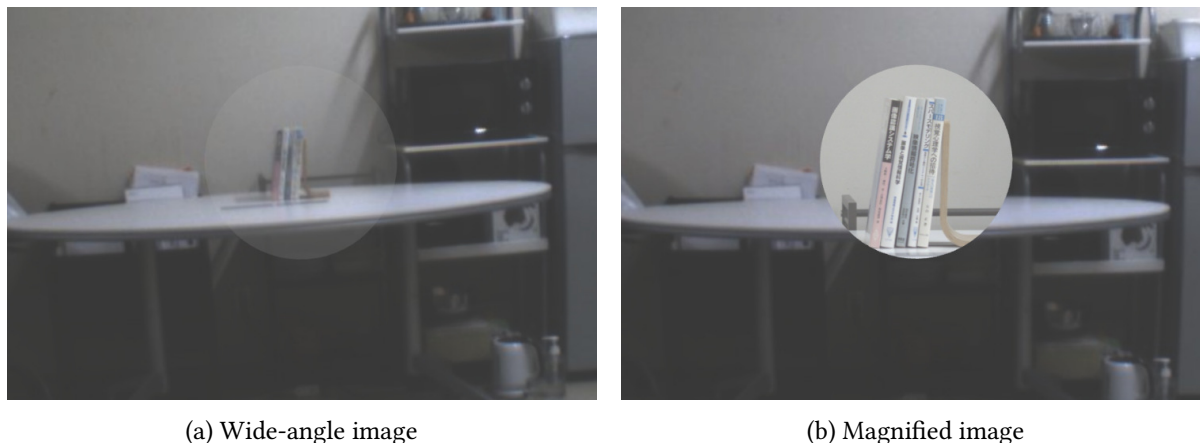


Figure 4: Switching image display. The magnified view overlays the wide-angle context when enabled.

3. Experiments and results

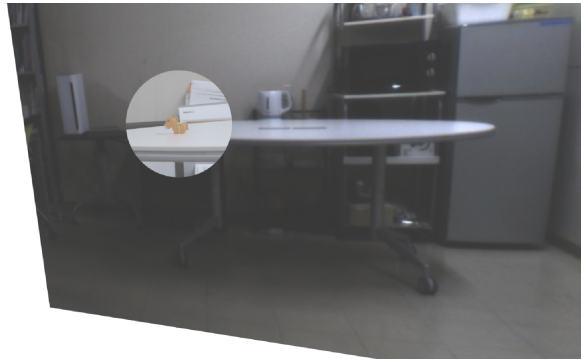
We conducted a qualitative evaluation of the prototype in an indoor environment to assess its performance and user experience. Figure 5 shows the HMD view and a user operating the system.

Our observations confirmed that the low-latency link between HMD motion and the ultrafast pan-tilt camera enabled smooth and immediate gaze shifts. Users could rapidly acquire magnified details at their point of regard without perceptible delay. The user-triggered display mechanism was reported to be a key factor for comfort. By giving users explicit control, any residual mismatch between head motion and the magnified image motion became significantly less distracting. This allowed for comfortable operation over extended periods, as users were not forced into a continuous, tight coupling with the high-magnification view. Because the magnified view was always presented at the center of the field of view, peripheral distortions were not a salient issue. We did observe slight misalignments between the wide-angle and magnified views when users looked toward the edges of the wide-angle image. While calibration successfully reduced this effect, perfect alignment remains a challenge due to factors like lens distortion and mechanical tolerances.

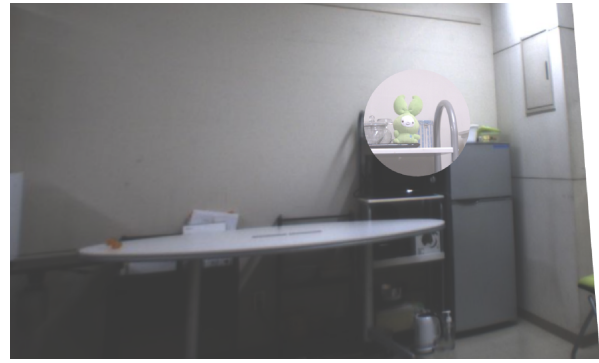
4. Conclusion and future work

We presented a VR remote magnified viewing system based on the principle of Optical Foveation. By using an ultrafast pan-tilt camera with a galvano mirror, our system delivers high-resolution details at the user’s gaze point with minimal latency, while a wide-field camera preserves the broader context. This approach, combined with a user-triggered display policy, effectively mitigates the discomfort typically associated with magnified views in VR, enabling rapid viewpoint control and comfortable remote inspection.

Future work will focus on two main areas. First, we plan to extend the system to support multiple concurrent users. The high-speed capability of the pan-tilt camera allows for time-multiplexing several magnified viewpoints, enabling different users to inspect distinct regions of interest within the same shared scene. A key challenge will be to manage the per-user refresh rate to maintain a comfortable experience. Second, acknowledging the limitations of our preliminary qualitative study, we will conduct



(a) Presented image (facing left)



(b) Presented image (facing right)



(c) User using the system (facing left)



(d) User using the system (facing right)

Figure 5: Experimental use case: user observing video from a remote location while switching between context and magnified detail.

more rigorous, quantitative evaluations. This will include measuring end-to-end latency from head motion to display update and performing controlled user studies. These studies will compare our system against conventional baselines, such as mechanical PTZ systems and digital zoom, to assess task performance, usability, and the reduction in VR sickness using standardized metrics like the Simulator Sickness Questionnaire (SSQ) [7].

Declaration on Generative AI

During the preparation of this work, the author used GPT-5 and gemini in order to: Grammar and spelling check. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

- [1] Aykut, Tamay, Lochbrunner, Stefan, Karimi, Mojtaba, Cizmeci, Burak, & Steinbach, Eckehard. (2017). A stereoscopic vision system with delay compensation for 360 remote reality. In *Proceedings of the on Thematic Workshops of ACM Multimedia 2017* (pp. 201-209).
- [2] Buhler, Helmut, Misztal, Sebastian, & Schild, Jonas. (2018). Reducing vr sickness through peripheral visual effects. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 517-519). IEEE.
- [3] Chang, Huiwen, & Cohen, Michael F. (2017). Panning and zooming high-resolution panoramas in virtual reality devices. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (pp. 279-288).
- [4] Chang, Eunhee, Kim, Hyun Taek, & Yoo, Byounghyun. (2020). Virtual reality sickness: a review of causes and measurements. *International Journal of Human-Computer Interaction*, 36(17), 1658-1682.

- [5] Farsiu, Sina, Robinson, M. Dirk, Elad, Michael, & Milanfar, Peyman. (2004). Fast and robust multiframe super resolution. *IEEE transactions on image processing*, 13(10), 1327-1344.
- [6] Guenter, Brian, Finch, Mark, Drucker, Steven, Tan, Desney, & Snyder, John. (2012). Foveated 3D graphics. *ACM Transactions on Graphics (TOG)*, 31(6), 1-10.
- [7] Kennedy, Robert S., Lane, Norman E., Berbaum, Kevin S., & Lilienthal, Michael G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3), 203-220.
- [8] Kim, Juno, Charbel-Salloum, Andrew, Perry, Stuart, & Palmisano, Stephen. (2022). Effects of display lag on vection and presence in the Oculus Rift HMD. *Virtual Reality*, 1-12.
- [9] Neves, Joao C., Moreno, Juan C., Barra, Silvio, & Proença, Hugo. (2015). Acquiring High-Resolution Face Images in Outdoor Environments: A Master-Slave Calibration Algorithm. In *Proceedings of 2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)* (pp. 1-8).
- [10] Ren, Yi, & Fuchs, Henry. (2016). Faster feedback for remote scene viewing with pan-tilt stereo camera. In *Proceedings of 2016 IEEE Virtual Reality (VR)* (pp. 273-274). IEEE.
- [11] Saharia, Chitwan, Ho, Jonathan, Chan, William, Salimans, Tim, Fleet, David J., & Norouzi, Mohammad. (2022). Image super-resolution via iterative refinement. *IEEE transactions on pattern analysis and machine intelligence*, 45(4), 4713-4726.
- [12] Stauffert, Jan-Philipp, Niebling, Florian, & Latoschik, Marc Erich. (2020). A survey on latency and its influence on presence and cybersickness. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 724-727). IEEE.
- [13] Xu, Yiliang, & Song, Dezhen. (2010). Systems and algorithms for autonomous and scalable crowd surveillance using robotic PTZ cameras assisted by a wide-angle camera. *Autonomous Robots*, 29(1), 53-66.