Enhancing Group Telepresence Interaction through Seamless Inverse Panorama Projection on Spherical Display

Nozomi Hayashida
linda@ucl.nuee.nagoya-u.ac.jp
Nagoya University
Nagoya, Japan

Shin Katayama
shinsan@ucl.nuee.nagoya-u.ac.jp
Nagoya University
Nagoya, Japan

Kenta Urano
urano@nagoya-u.jp
Nagoya University
Nagoya, Japan

Takuro Yonezawa
takuro@nagoya-u.jp
Nagoya University
Nagoya, Japan

Nobuo Kawaguchi
kawaguti@nagoya-u.jp
Nagoya University
Nagoya, Japan

Figure 1: Demonstration of MetaPo's spherical display system in action. The image shows participants interacting with the spherical display, which projects a 3D character. This system allows users to experience an immersive and interactive virtual environment, enhancing remote communication and collaboration.

Abstract

In recent years, advancements in telepresence technology have made remote communication and collaboration easier. MetaPo is a system that uses spherical displays to seamlessly integrate physical and digital interactions. However, conventional inverse panorama techniques face challenges such as image seams, coverage issues, and overlapping displays. To address these challenges, this study proposes an inverse panorama technique based on coordinate transformation, enhancing visual continuity and user experience. Additionally, evaluation methods necessary to demonstrate the effectiveness of the proposed method in improving the quality of remote communication have been devised. This research provides a novel solution to challenges in spherical display technology and paves the way for more immersive telepresence experiences.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
UbiComp Companion '24, October 5–9, 2024, Melbourne, VIC, Australia.
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-1058-2/24/10
https://doi.org/10.1145/3675094.3678369

CCS Concepts

• **Human-centered computing** → **Virtual reality**; **Collaborative and social computing**.

Keywords

Inverse Panorama, Spherical Display, Telepresence Robot, Remote Communication, Virtual Reality

ACM Reference Format:
Nozomi Hayashida, Shin Katayama, Kenta Urano, Takuro Yonezawa, and Nobuo Kawaguchi. 2024. Enhancing Group Telepresence Interaction through Seamless Inverse Panorama Projection on Spherical Display. In *Companion of the 2024 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '24), October 5–9, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3675094.3678369

1 Problem Statement

Telepresence technologies are rapidly advancing to enhance remote communication. Notable examples include Cisco's TelePresence system[1] and Microsoft Teams Rooms[8], which enable participants in disparate locations to conduct meetings as if co-located. Furthermore, VR and AR technologies such as Meta's Quest and Microsoft's HoloLens provide new platforms for interaction and remote collaboration in virtual environments[5, 9].

This research aims to achieve seamless communication across both physical and virtual spaces, with a particular emphasis on addressing the asymmetries that arise between these environments. In physical spaces, participants share the same environment and receive natural visual, auditory, and tactile feedback, enabling intuitive communication. Conversely, in virtual spaces, participants are separated by distance, often resulting in limited feedback, thus affecting the quality of communication.

These asymmetries are particularly evident when gestures and eye movements made by participants in the physical space are not accurately conveyed to those in the virtual space, or when the viewpoints of virtual participants do not align with the physical context. Such discrepancies can impede mutual understanding and collaboration.

To address these challenges, Yonezawa et al. developed the MetaPo group telepresence robot, which is equipped with a spherical display[11, 12, 14]. This system utilizes two connection modes, Mixed Link and Immersive Link, to facilitate seamless interaction between physical and digital realms(Figure 2). In Mixed Link mode, users interact with a central 360-degree camera, and the panoramic video displays on a remote spherical display. This setup provides a broader field of view than traditional video calls, enabling more comprehensive and natural interactions with remote participants.

In Immersive Link mode, users virtually enter a dome-shaped space (portal) within MetaPo, offering a more immersive experience. This method is analogous to single-user VR telepresence but provides a group telepresence experience with panoramic and mobility features. In this mode, as users within the portal interact while viewing external video, inverse panorama projection techniques are required to project the portal space onto the spherical display. Previous studies have documented the use of spherical displays in exhibition and entertainment contexts[10, 13]. However, these technologies are rarely employed in the context of communication, resulting in limited literature on inverse panorama techniques.

Our previous research involved positioning multiple cameras and stitching the images together, but this method presents several challenges:

- **Image Seams**: When stitching images from multiple cameras, noticeable seams can occur, especially with moving objects or complex backgrounds, disrupting visual continuity.
- **Increased Computational Cost**: The image stitching process requires advanced computational processing, making real-time performance difficult and potentially causing system delays.
- **Coverage and Overlapping Display Issues**: Depending on the number and arrangement of cameras, certain areas within the portal may not be fully covered, or the same user or object may be captured multiple times from different angles, leading to visual inconsistencies and confusion.

We propose an inverse panorama technique using coordinate transformation to address these challenges. This method aims to enhance visual continuity and user experience in MetaPo's Immersive Link.

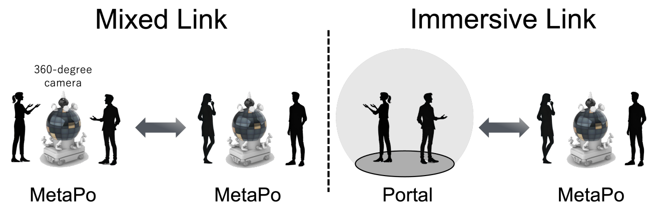

Figure 2: Illustration of Mixed Link and Immersive Link modes in MetaPo's spherical display system. The left side shows the Mixed Link mode, where remote participants appear as if they are in the same space. The right side shows the Immersive Link mode, where users enter a virtual dome-shaped space for an immersive communication experience.

2 Related Work

2.1 Spherical Displays

Spherical displays significantly enhance communication and interaction by providing 360-degree content, surpassing traditional flat displays through immersive experiences. Williamson et al. [13] deployed the spherical display application "GlobalFestival" at an international music festival over four days, evaluating two types of touch interaction techniques. Englmeier et al. [2] proposed a novel method using a physical sphere as a tactile proxy for manipulating virtual objects in an AR environment.

Li et al. [7] proposed OmniGlobeVR, a 360-degree communication system using a spherical display to enable virtual reality interactions. This system allows users to manipulate 360-degree video content within a shared virtual environment, enhancing the sense of presence and collaboration among participants. However, their study also highlighted the challenges of maintaining visual continuity and synchronization among multiple users.

Overall, while spherical displays offer promising opportunities for enhancing communication and interaction, further research is needed to address the challenges associated with applications focused on collaboration and communication.

2.2 Inverse Panorama Technology

Inverse panorama technology is essential for providing a seamless 360-degree visual experience. This technology typically integrates multiple camera feeds into a single continuous panorama, offering users a consistent and immersive visual environment. Gross et al. [3] developed a system that acquires real-time 3D video from multiple cameras, creating photorealistic 3D inlays of users. This method demonstrates the feasibility of high-quality 3D video capture but highlights the significant complexity involved in reconstructing entire spatial environments.

The proposed method transforms Cartesian coordinates into polar coordinates, allowing seamless projection of single-camera content onto spherical displays. This approach addresses several inherent challenges in traditional multi-camera setups, such as image seams, computational cost, and visual consistency.

3 Methodology

In this study, we propose a inverse panorama technique using coordinate transformation to enhance visual continuity and user experience in MetaPo's Immersive Link.

3.1 Challenges of Inverse Panorama Technique

The inverse panorama technique aims to project virtual environments onto a spherical display. However, integrating images from multiple cameras requires complex processing, which can result in seams and misalignments within the images. Additionally, depending on the number and placement of cameras, some areas within the portal may not be fully covered. Multiple cameras capturing the same user or object from different angles can lead to visual inconsistencies and confusion when the same subject is displayed multiple times (Figure 3a). These limitations pose significant barriers to achieving an immersive and interactive communication paradigm, thereby limiting the overall effectiveness of the technology and user engagement.

3.2 Proposed Inverse Panorama Technique

The inverse panorama method proposed in this study innovates spatial representation by transforming portal coordinates from Cartesian coordinates (x, y) to polar coordinates (r, θ). This technique enables the capture of transformed images using a single camera (Figure 3b). In this method, objects are positioned based on their angle and distance relative to the center of the spherical display, integrating omnidirectional views into a single image.

Specifically, the coordinate transformation is performed using the following equations:

$$r = \sqrt{x^2 + y^2}$$

$$\theta = \arctan\left(\frac{y}{x}\right)$$

Adopting this polar coordinate framework allows for more natural projection of images onto the curved surface of the spherical display, which is challenging to achieve with the traditional Cartesian coordinate system. This enables harmonious integration of scenes from various perspectives, maintaining visual consistency even during viewpoint changes on the spherical display.

3.3 System Configuration

The overall configuration of the proposed system is as follows:

(1) **Portal**: A dome-shaped virtual space within MetaPo.
(2) **Transformation Space**: The space after transforming the coordinates within the portal to polar coordinates.
(3) **Camera**: An orthographic camera is placed to capture the transformed space.

This system configuration allows the proposed inverse panorama technique to improve visual continuity and user experience, realizing an immersive remote communication in MetaPo's Immersive Link.

4 Evaluation

To verify the effectiveness of the proposed inverse panorama technique using coordinate transformation, both quantitative and qualitative evaluations are essential. These evaluations will compare the proposed method with conventional multi-camera stitching methods (using 3 and 6 cameras), focusing on computational cost, completeness of field of view, and the quality of communication.

To objectively measure the improvements brought by the proposed method in terms of computational cost, the stitching costs associated with multiple cameras will be compared to the cost of coordinate transformation in the proposed method. Metrics such as average processing time, maximum processing time, CPU usage, GPU usage, and memory usage will be included. Regarding the completeness of the field of view, the coverage differences provided by different virtual camera arrangements will be evaluated. This includes comparing the proposed method with conventional methods using 3 and 6 cameras. Evaluation metrics will include the calculation of the coverage rate, which is the proportion of the portal fully covered, and the detection of blind spots by measuring the number and area of blind spots in the images for each method. The number of blind spots, average blind spot area, and maximum blind spot area will also be considered.

To assess user experience and the quality of communication, user feedback will be gathered through surveys and interviews. To evaluate user immersion, presence, ease of communication, and cognitive load, the ITC-Sense of Presence Inventory (ITC-SOPI) [6] and NASA-TLX [4] will be used. Interviews will focus on gaining deeper insights into users' experiences and preferences, with a focus on visual continuity, perceived seams or artifacts, and the impact on communication.

These evaluations are crucial to confirm the practicality and superiority of the proposed method. By clarifying the extent to which the proposed method resolves the issues of conventional methods in terms of computational cost, completeness of field of view, and the quality of communication, we can demonstrate the effectiveness of the proposed method and establish a foundation for further development.

The ultimate goal of this study is to significantly enhance visual continuity and user experience in MetaPo's Immersive Link, thereby improving the quality of remote communication. Based on the evaluation results, it is expected to confirm the practicality of the proposed method and offer a new approach to improving the quality of remote communication.

5 Contribution

This study aims to enhance visual continuity and user experience in MetaPo's Immersive Link by developing an inverse panorama technique using coordinate transformation, ultimately improving the quality of remote communication.

Firstly, an inverse panorama technique that leverages the transformation from Cartesian coordinates to polar coordinates is proposed. This technique enables seamless 360-degree image generation with a single camera, effectively addressing the issues of seams and misalignments commonly found in conventional multi-camera stitching methods.

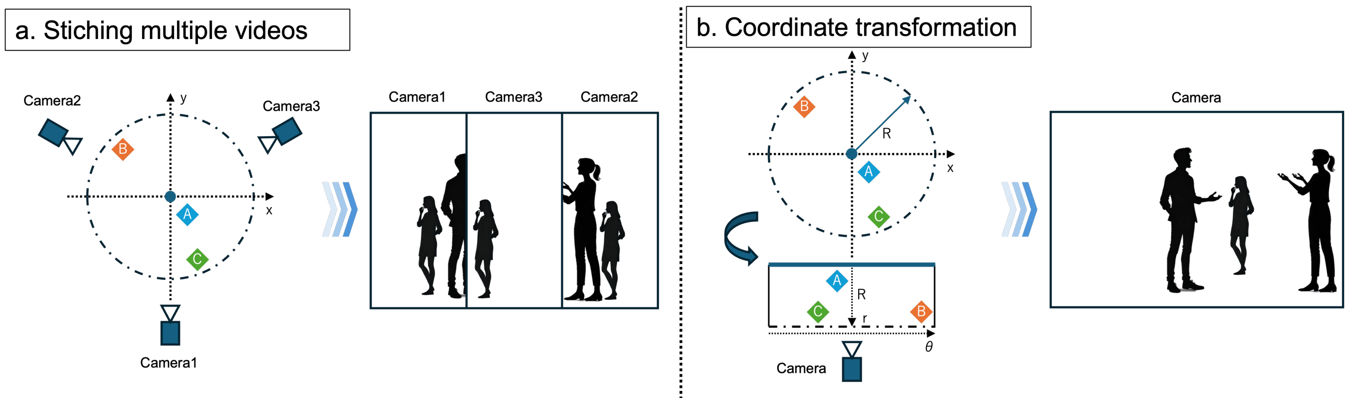

Figure 3: Illustration of the implementation of the inverse panorama technique. The left part of the figure shows multiple cameras (1, 2, 3) positioned to capture images from different viewpoints. The conventional multi-camera stitching method integrates these camera feeds, which can lead to breaks and geometric inconsistencies. The right part of the figure demonstrates the proposed inverse panorama technique, where Cartesian coordinates are transformed into polar coordinates, resulting in a seamless and consistent panoramic image.

Secondly, a comparative evaluation between the proposed method and traditional multi-camera stitching techniques is conducted. This evaluation elucidates the respective advantages and disadvantages, serving as a reference for future research and providing guidelines for other researchers developing new methodologies.

Finally, the impact of the inverse panorama technique on user experience is investigated through user surveys and quantitative evaluations. These insights are invaluable for designing more effective telepresence systems, offering a comprehensive understanding of how the proposed technique influences user perception and interaction.

Acknowledgements.

This work was supported by JST CREST Grant Number JPMJCR22M4, JST RISTEX Grant Number JPMJRS23K3 and Council for Science, Technology and Innovation, "Cross-ministerial Strategic Innovation Promotion Program (SIP), Development of foundational technologies and rules for expansion of the virtual economy"(JPJ012495).(funding agency: NEDO).

References

[1] Cisco Systems, Inc. 2023. Cisco TelePresence Systems. https://www.cisco.com/c/en/us/products/collaboration-endpoints/immersive-telePresence/index.html. Accessed: 2023-03-10.

[2] David Englmeier, Julia Dörner, Andreas Butz, and Tobias Höllerer. 2020. A Tangible Spherical Proxy for Object Manipulation in Augmented Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 221–229. https://doi.org/10.1109/VR46266.2020.00041

[3] Markus Gross, Stephan Würmlin, Martin Naef, Edouard Lamboray, Christian Spagno, Andreas Kunz, Esther Koller-Meier, Tomas Svoboda, Luc Van Gool, Silke Lang, et al. 2003. blue-c: a spatially immersive display and 3d video portal for telepresence. *ACM Transactions on Graphics (TOG)* 22, 3 (2003), 819–827.

[4] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[5] Nozomi Hayashida, Hironori Shimosato, Kenta Urano, Takuro Yonezawa, and Nobuo Kawaguchi. 2023. Mobility Link XR: Interspace Interaction System in Electric Wheelchair. In *International Conference on Human-Computer Interaction*. Springer, 314–320.

[6] Jane Lessiter, Jonathan Freeman, Edmund Keogh, and Jules Davidoff. 2001. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments* 10, 3 (06 2001), 282–297. https://doi.org/10.1162/105474601300343612 arXiv:https://direct.mit.edu/pvar/article-pdf/10/3/282/1623678/105474601300343612.pdf

[7] Zhengqing Li, Liwei Chan, Theophilus Teo, and Hideki Koike. 2020. OmniGlobeVR: A collaborative 360 communication system for VR. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[8] Microsoft. 2023. Microsoft Teams Rooms. https://www.microsoft.com/en-us/microsoft-teams/microsoft-teams-rooms. Accessed: 2024-01-26.

[9] Hironori Shimosato, Nozomi Hayashida, Kenta Urano, Takuro Yonezawa, and Nobuo Kawaguchi. 2022. XR Communication System for Remote Control Wheelchairs. In *Proceedings of the 12th International Conference on the Internet of Things*. 219–223.

[10] Nikita Soni, Schuyler Gleaves, Hannah Neff, Sarah Morrison-Smith, Shaghayegh Esmaeili, Ian Mayne, Sayli Bapat, Carrie Schuman, Kathryn A Stofer, and Lisa Anthony. 2020. Adults' and children's mental models for gestural interactions with interactive spherical displays. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[11] Yoshiki Watanabe, Nozomi Hayashida, Shin Katayama, Kenta Urano, Takuro Yonezawa, and Nobuo Kawaguchi. 2023. 360 Robot Hand Interaction for Group Telepresence with MetaPo. In *Proceedings of the 13th International Conference on the Internet of Things*. 183–185.

[12] Yoshiki Watanabe, Nozomi Hayashida, Shin Katayama, Kenta Urano, Takuro Yonezawa, and Nobuo Kawaguchi. 2023. 360 Robot Hand System for Omnidirectional Interaction by Multiple Users. In *Proceedings of the 13th International Conference on the Internet of Things*. 295–299.

[13] Julie R Williamson, Daniel Sundén, and Jay Bradley. 2015. GlobalFestival: evaluating real world interaction on a spherical display. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 1251–1261.

[14] Takuro Yonezawa, Nozomi Hayashida, Kenta Urano, Johannes Przybilla, Yutaro Kyono, and Nobuo Kawaguchi. 2022. MetaPo: A Robotic Meta Portal for Interspace Communication. In *ACM SIGGRAPH 2022 Posters*. 1–2.