Theory of Spherical VR model for Landscape Representation

Hiroyuki Nishimoto*, Toru Takahashi† and Masakazu Yoshida‡
* Osaka Sangyo University, Japan
E-mail: nishimoto@ise.osaka-sandai.ac.jp
† E-mail: takahashi@ise.osaka-sandai.ac.jp
‡ E-mail: yoshida@ise.osaka-sandai.ac.jp

***Abstract*—Spherical images capture visual information from all directions in a single frame, making them an essential tool for recording and analyzing landscapes. However, these images differ significantly from conventional photographs taken by cameras or from human visual perception, often resulting in a sense of visual discomfort. To address this issue, this study constructs the theory of a mathematical model for geometrically describing and analyzing landscape representations in spherical images. In particular, the model captures angular deviations that arise from differences in depth to the observed objects as the viewpoint moves. The proposed model is defined by a geometric structure that includes points at infinity in three-dimensional directions, distinguishing it from the conventional Riemann sphere.**

I. Introduction

Typically, virtual reality (VR) space is represented in a 2D planar form realized by equirectangular projection [**1-3**]. However, in a VR space, the user wears a head-mounted display and moves their head and body freely, causing the viewpoint to change continuously and nonlinearly [**4-8**]. For this reason, in equirectangular projection, the correction of panoramic distortion that accompanies viewpoint movement and the consistency of depth and parallax become issues [**9-12**]. This suggests that a numerically feasible algorithm is needed to deal with the inconsistencies that arise when projecting a perspective space, in which a vanishing point is defined at infinity [**13**], into a finite rendering space [**1**].

While prior research has explored various methods for capturing and visualizing 360-degree imagery, many approaches rely on conventional spherical projections such as the Riemann sphere, which often fail to account for scale distortions and depth-related perceptual inconsistencies, especially when the viewpoint changes. These limitations can lead to visual discomfort or misinterpretation when omnidirectional content is applied to immersive environments such as VR [**14-15**].

To address these challenges, we use a previously developed visual space model [**16**] to construct a theory of a VR space model for landscape representation. Fig. 1 shows that by using the visual space model as a projective space, the mapping from perspective space to real space can be realized as a simple model. Perspective space refers to a perceptual and visual spatial structure based on linear perspective [**17**]. It represents how humans perceive depth and spatial relationships when viewing the three-dimensional world as a two-dimensional visual field, such as in images or paintings. A perspective space is a space based on projection geometry, and is defined by a linear projection from the viewpoint (the camera origin).

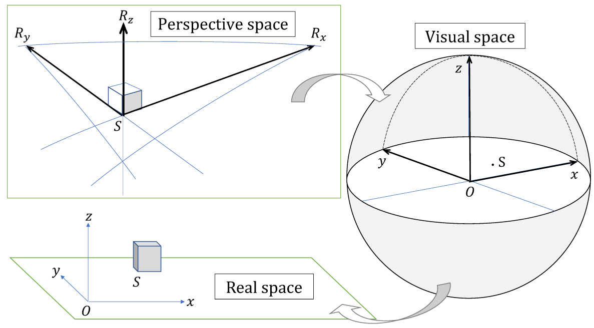

Fig. 1. Complex projective space.

The proposed model formulates a unified mathematical framework in the form of a differential equation that encompass multiple conditions related to scale in visual space. In particular, it mathematically models the angular deviation caused by the depth difference between the viewpoint and an object as the viewpoint moves. Furthermore, the model makes it possible to calculate the scale of objects, enabling the generation of landscape images with reduced perceptual distortion in VR environments. This approach provides a new geometric foundation for more accurate and intuitive landscape representation in spherical imaging.

In Sections 2, we describe the core theory of the spherical VR model for landscape representation. Specifically, we show that the proposed model can be expressed in the form of a differential equation that satisfies four conditions related to scale, which are scale observability, the inverse square law, linear perspective and cosine depth in a unit sphere.

In Section 3, we show that the perspective depth can be expressed as a geometric series, which allows us to calculate the scale of the object. And we derive the perceived depth to an object in order to accurately represent the perceptual changes caused by the depth due to viewpoint movement in landscape representation. Specifically, we show that the perceived depth is a relative value of the radius of the visual space. This leads to the conclusion that the lens magnification of the head-mounted display affects the radius of the VR space. In Section 4, we formulate landscape representation in spherical images when

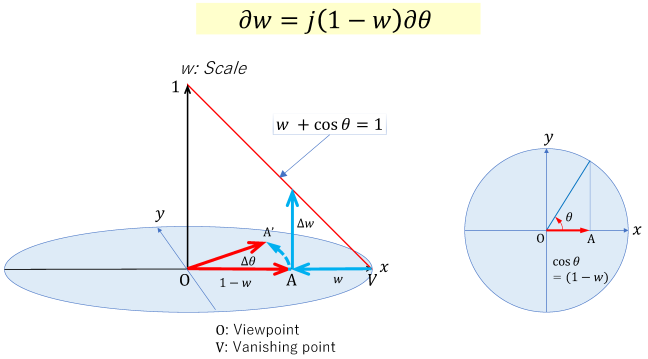

Fig. 2. Differential equation for the observability of scale [**16**].

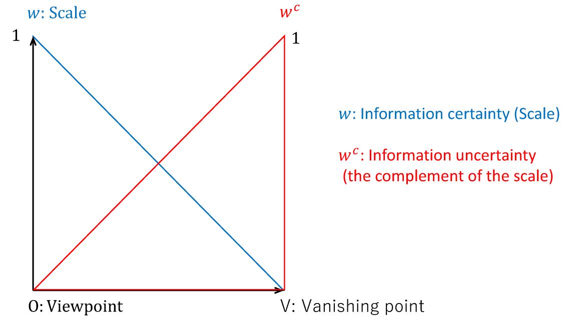

Fig. 3. Scale as information certainty, the complement of the scale as information uncertainty.

the viewpoint moves. In Section 5, we discuss the equidistant and orthogonal projections for this optical system. Section 6 summarizes the theory behind constructing a spherical VR space.

II. A Visual Space Model That Connects Perspective Space and Real Space

In this section, as shown in Fig. 2, we propose a visual space model for geometrically describing and analyzing landscape representation in spherical images [**16**]. First, we show that the proposed model can be expressed in the form of a differential equation that satisfies four scale-related terms.

The proposed visual space model considers the z-component of scale w in which the viewpoint is the origin, the line of sight is on the x-axis, and the deflection angle θ with respect to the x-axis is on the x-y plane. The imaginary number j, taking into account the quaternion, represents a rotation of 90 degrees from the y-axis to the z-axis. In this model, the scale of the origin O is set to 1.0 as the maximum probability and the differential equation for the observability of scale is given by

$$\partial w = j\,(1 - w)\,\partial\theta \quad \left(-\frac{\pi}{2} < \theta < \frac{\pi}{2} \right). \tag{1}$$

In this model, the following scale-related terms are combined into one differential equation.

1. Scale observability
2. Inverse square law
3. Linear perspective
4. Cosine depth in a unit sphere

The following sections explain each term of the differential equation.

A. *Scale observability*

As for the observability of scale, there is a characteristic that the scale and perspective of a point moving along the line of sight cannot be detected. This is because the moving direction of the point and the direction of the line of sight are the same. In other words, it is because the deflection angle between the point's direction movement and the line of sight is zero. For example, by moving the viewpoint, the deflection angle to the object changes, making the scale observable. If we consider binocular disparity as the instantaneous movement from the right eye to the left eye, it is also represented by deflection angle. Therefore, in this model, the change in scale Δw is explained by the change in the deflection angle $\Delta\theta$. In this way, the observability of scale is expressed by the cross product of the line of sight deviation.

B. *Inverse square law*

Since the imaginary number j represents a rotation of 90 degrees, (1) shows that the scale change on the left side is perpendicular to the scale change on the right side. Therefore, the imaginary number j represents the inverse square law, indicating that the same change in scale is also observed perpendicular to the observed change in scale. Similarly, the imaginary number j means that a change in scale in the vertical direction relative to the line of sight is observed due to a change in depth in the line of sight direction. On the other hand, the inverse square law also assumes that the vertical and horizontal scale changes are equal, so the imaginary number j in the formula (1) means the inverse square law.

C. *Linear perspective*

Integrating (1) using Euler's formula, the solution is given by

$$w = 1 - \cos\theta + j\sin\theta \quad \left(-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2} \right). \tag{2}$$

The real and imaginary parts represent the scale changes in the x-axis and z-axis, respectively. As shown in Fig. 2, the slope of the scale between O and V means linear perspective in visual space. At the vanishing point V, the scale of all objects becomes zero and they are no longer visible, so it is the boundary of visual space. Thus, according to linear perspective, the distance to the boundary of visual space is considered equal in all directions, defining a radius and thus forming visual space as a sphere.

As shown in Fig. 3, when scale w is defined as the probability of information certainty, then perspective depth [**16**]

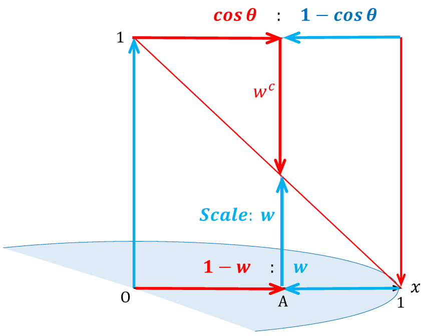

Fig. 4. Perspective depth is expressed by the complement of the cosine θ.

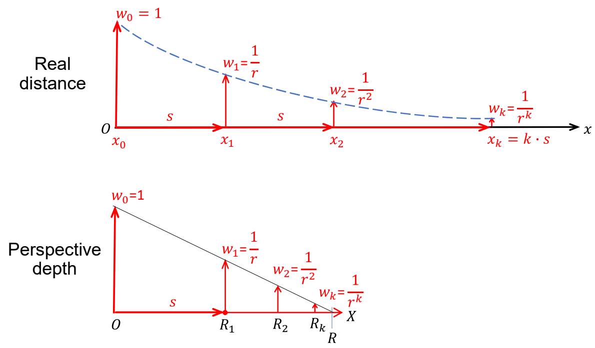

Fig. 5. Real distance vs. perspective depth [**16**]

can be defined as the probability of information uncertainty as the complement of scale w^c. Therefore, the condition for the probability density function is given by

$$w + w^c = 1 \quad (0 \leq w \leq 1, \quad 0 \leq w^c \leq 1)\,. \tag{3}$$

Since the scale at the origin is 1.0, visual space can be defined as a unit sphere. Since the radius of the unit sphere is 1.0, $\cos\theta$ can express the perspective depth from the viewpoint to the vanishing point as the cosine depth. Similarly, $\sin\theta$ can freely express the change in scale from the viewpoint to the vanishing point. Therefore, (1) shows all points within the sphere can be represented by deflection angle θ.

D. Cosine depth in a unit sphere

Since w^c is the perspective depth and it can be expressed by $\cos\theta$, the above equation is also given by

$$w + \cos\theta = 1 \quad \left(0 \leq w \leq 1, \quad 0 \leq \theta \leq \frac{\pi}{2}\right). \tag{4}$$

As shown in Fig. 4, the condition for linear perspective is also given by the following equation:

$$\sin\theta + \cos\theta = 1 \quad \left(0 \leq \theta \leq \frac{\pi}{2}\right). \tag{5}$$

Therefore, under the condition that the sum of scale and perspective depth is 1, the above probability density function can be expressed as a square with 1 side as shown in Fig. 3.

Moreover, as shown in Fig. 2, (1) describes a change in scale along the z-axis, which corresponds to an imaginary dimension, induced by angular deviations occurring in the real two-dimensional x-y plane. By extending this model to the y-z and z-x planes, it can be generalized to a complex three-dimensional (i.e., real six-dimensional) spatial model.

In this way, in calculations related to visual space, (5) appears in which the sum of scale and perspective depth is 1 appears. This suggests that linear perspective and elliptical space are closely related. This also suggests that the linear perspective condition can be used instead of elliptic equations when calculating positions in visual space.

III. Perspective depth in visual space

In this section, we derive the perspective depth that satisfies the perspective condition in (5), which is necessary for landscape representation in spherical images.

Fig. 5 demonstrates the relationship between the real distance in $\mathbb{R}^3$ as physical space and the perspective depth [**18**] in visual space.In the upper part, let s be the real distance from point O where the object appears to be $1/r$ times larger. If the real distance is twice s, the object will appear $1/r^2$ times larger. The lower part demonstrates the relationship between scale and depth from the viewpoint, the origin of visual space. This perspective depth is defined as the real distance multiplied by $1/r$ when the distance is extended by s according to linear perspective.

As shown in Fig. 5, when k is a multiple of the unit distance s, x_k is defined as follows:

$$x_k = k \cdot s \quad (k > 0, \quad s > 0)\,. \tag{6}$$

In real space $\mathbb{R}^3$, the scale w_k on the point x_k is expressed by the following equation, where $1/r$ is the depth reduction rate at unit distance s:

$$w_k = \frac{1}{r^k} \quad (k > 0, \quad r > 1)\,. \tag{7}$$

From (7), perspective depth R_k is given by the parameters as demonstrated in Fig. 1:

$$R_k = s \sum_{n=0}^{k-1} \frac{1}{r^n} = \frac{s}{1 - \dfrac{1}{r}} \left(1 - \frac{1}{r^k}\right). \tag{8}$$

Therefore, the perspective depth R from the viewpoint to the vanishing point is given by

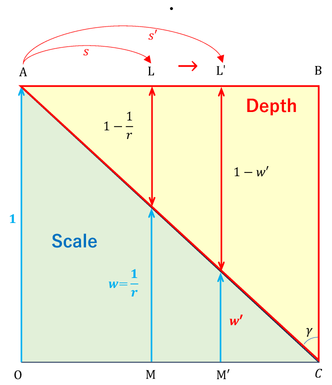

Fig. 6. Perceived depth is relative to the radius of visual space.

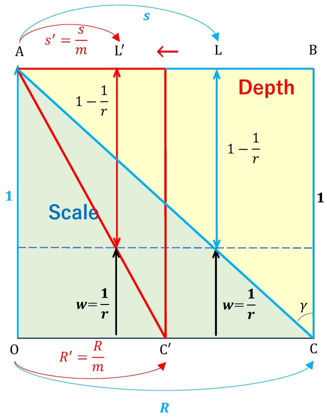

Fig. 7. Lens magnification reduces the radius of the visual space.

$$R = \lim_{k\to\infty} \frac{s}{1-\frac{1}{r}}\left(1-\frac{1}{r^k}\right) = \frac{s}{1-\frac{1}{r}}. \tag{9}$$

In this case, the arbitrary perspective depth R_k is given by

$$R_k = R\left(1-\frac{1}{r^k}\right). \tag{10}$$

IV. Landscape Representation in Spherical Images when the Viewpoint is Moved

In this section, we formulate landscape representation in spherical images when the viewpoint moves.

A. Perceived depth is relative to the radius of visual space

First, we derive the perceived depth to objects in order to accurately represent the perceptual changes in depth caused by viewpoint movement.

To derive perspective depth through drawing, as shown in Fig. 6, draw △OAC as a change in scale, and then draw △ABC on top of it as the condition for linear perspective, since the sum of scale and perspective depth is 1. MN denotes the depth reduction ratio $1/r$ at point L, and LM denotes the perspective depth $(1-1/r)$. Similarly, M'N' indicates the depth reduction ratio $1/r'$ at point L', and L'M' indicates the perspective depth $(1-1/r')$. When ∠BCA is γ, the radii of the spheres at points L and L' can be expressed relatively as follows:

$$\begin{aligned} R &= \frac{s}{1-\frac{1}{r}} = \tan\gamma \quad \left(0<\gamma<\frac{\pi}{2}\right). \\ R' &= \frac{s'}{1-\frac{1}{r'}} = \tan\gamma \quad \left(0<\gamma<\frac{\pi}{2}\right). \\ \therefore R &= R'. \end{aligned} \tag{11}$$

Therefore, the radius R of the visual space sphere is constant regardless of the depth reduction rate $1/r$ or the depth s.

To derive the perspective depth through a lens by drawing, as shown in Fig. 7, △OAC is drawn to represent the change in scale, and then △ABC is drawn on top of it as the condition for linear perspective, since the sum of the scale and the perspective depth is 1. If the depth when the scale is $1/r$ is s and the depth when the same scale is viewed through a lens is s', then the lens magnification m times is given by the following formula.

$$s' = \frac{s}{m} \quad (m>0). \tag{12}$$

Therefore, the radius of the sphere R' when the lens magnification is m is given by the following equation.

$$R' = \frac{s'}{1-\frac{1}{r}} = \frac{s}{1-\frac{1}{r}}\cdot\frac{1}{m} = \frac{R}{m}. \tag{13}$$

Therefore, when the lens magnification is m, the radius of the visual space sphere becomes $1/m$. When the lens magnification is m, the radius R' is expressed in γ' as follows:

$$R' = \tan\gamma' = \frac{1}{m}\cdot\tan\gamma \quad \left(0<\gamma'<\frac{\pi}{2}\right) \tag{14}$$

This suggests that lens magnification affects perceived depth. This is because it can be mathematically explained that when the lens magnification is doubled, the radius of the visual space is halved, and the sense of depth to the vanishing point is also halved. In this model, (14) can be used as a reference to describe the cramped feeling when looking through a head-mounted display.

B. Parallel lines ride on ellipses with the same major axis in visual space

Parallel lines never intersect in physical space. By contrast, parallel lines intersect at the vanishing point in visual space. As shown in Fig. 8, if visual space is spherical, the condition for parallel lines to intersect at infinity can be shown by having

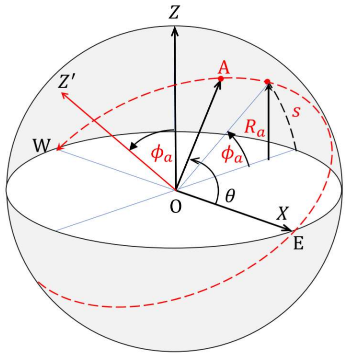

Fig. 8. Condition that parallel lines intersect at infinity [16].

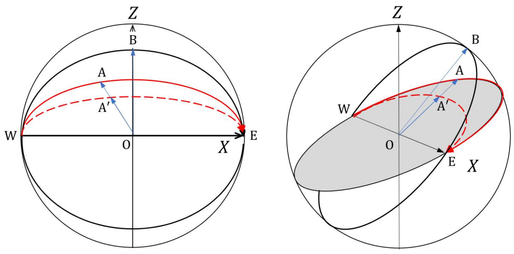

Fig. 9. Parallel lines ride on ellipses with the same major axis [16].

the dashed circle intersects at infinity points E and W on the X-axis. Because a straight line projected onto the spherical surface looks like an arc when viewed three-dimensionally. Thus, line WE and arc WAE appear to be parallel in visual space. As shown in Fig. 8, when viewpoint O, point A, and point A' are on the same line, arc WAE and arc WA'E overlap in visual space. Thus, arcs WAE and WA'E appear to be parallel in visual space. In addition, as arc WBE also appears to be parallel to line WE, arc WBE and arc WA'E appear to be parallel in visual space.

If the inclination angle of the dashed circle to the Z-axis is ϕ_a, the coordinates of A on the dashed circle in its orthogonal projection onto the X-Z plane are given by the following equation using the rotation angle θ of the Z'-axis:

$$A = (R\cos\theta, \quad R_a \sin\theta) \quad (R_a = R\sin\phi_a)\,. \tag{15}$$

As shown in Fig. 9, (15) represents an ellipse whose major axis is the sphere's diameter. A line through the origin can be part of the ellipse in visual space. Therefore, any straight line in $\mathbb{R}^3$ can be mapped on the ellipse in visual space. In other words, as parallel lines in physical space are transformed into ellipses with the same major axis in visual space, the direction of the major axis points the same direction as the parallel lines.

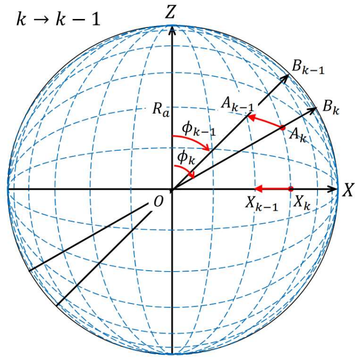

Fig. 10. Movement of points when the viewpoint is moved [16].

C. Movement of points when the viewpoint is moved

(10) and (15) are used to formulate the image change when the viewpoint is moved. In particular, we mathematically model the change in the deflection angle due to the difference in depth to the observed object that occurs as the viewpoint moves. Actual head-mounted displays use binocular disparity to represent the depth of an object. This requires the inter-pupillary distance to be adjusted for each individual. However, in this section, we consider a model that allows stereoscopic vision by projecting images onto a spherical surface in VR space, like a planetarium, and mathematically reproducing the difference in deviation angle due to depth differences.

As shown in Fig. 10, from (10), the coordinates of X_k are given by

$$X_k = (R_k, \quad 0) = \left(R\left(1 - \frac{1}{r^k}\right), \quad 0\right). \tag{16}$$

Further, the coordinates of X_{k-1} is given by

$$X_{k-1} = \left(R_k - \frac{s}{r^{k-1}}, 0\right). \tag{17}$$

To find the extreme values of each coordinate component, let $\phi = \pi/2 - \theta$ in (15). Then, the coordinates of point A_k are difined as

$$A_k = (R_k, \quad S_k) = (R\sin\phi_k, \quad R_a\cos\phi_k)\,. \tag{18}$$

As the solutions of (18), the coordinates of A_k are given by

$$A_k = \left(R\left(1 - \frac{1}{r^k}\right), \quad R_a\left(\frac{1}{r^k}\right)\right). \tag{19}$$

This makes it possible to simulate a chang in scenery in a spherical image as the viewpoint changes. This is because it can map the coordinates on the grid [19] to the coordinates in spherical geometry. Thus, the model shows that a point in real space can be mapped to the corresponding point in visual space. Therefore, this model is a mathematical model suitable for representing landscapes in spherical images using immersive environments such as virtual reality (VR).

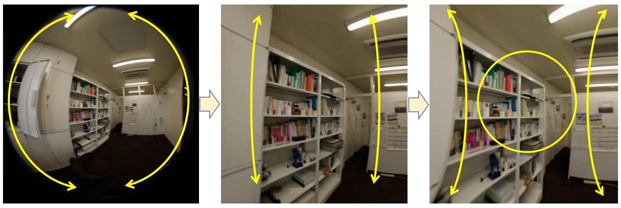

Fig. 11. Equidistant and orthogonal projections

V. Equidistant and orthogonal projections

Equidistant projection converts the central angle of the visual space to an equal interval [**20**]. For this reason, it seems that the most natural spherical VR space can be realized by applying equidistant projection. However, equidistant projection is not necessary for a spherical VR space to which the visual space model is applied. An orthogonal projection is basically sufficient for this optical system configuration.

The left side of Fig. 11 is an orthogonally projected image (fisheye image), the center is an equidistantly projected image, and the right is a doubly equidistantly projected image. Parallel lines projected onto the spherical surface of visual space become ellipses, but equidistant projection makes these ellipses closer to parallel lines. However, this is a discussion that applies when mapping a three-dimensional sphere onto a two-dimensional plane. When viewed through a head-mounted display, the ellipses in the proposed visual space model become parallel lines. For this reason, in a spherical VR model, assuming a pinhole lens, image processing using orthogonal projection rather than equidistant projection is sufficient.

VI. Conclusion

This study focuses on the characteristics of scale and proposes a novel geometric model for landscape representation in omnidirectional images. To address perceptual distortions caused by depth and viewpoint changes—issues that conventional spherical projections struggle to resolve—the proposed model introduces a unified mathematical framework based on differential equations. This framework enables the formalization of angular deviations resulting from viewpoint shifts and allows for accurate computation of the scale of observed objects. As a result, it offers the potential for intuitive and less distorted landscape representation even in VR environments.

It is a visual space model for geometrically describing and analyzing landscape representations in spherical images [**16**]. This model is based on projection geometry, and is defined by a linear projection from the viewpoint (the camera origin). It is a differential equation that incorporates the observability of scale, the inverse square law, and the linear perspective condition. The solution shows that visual space is a sphere based on elliptical geometry. Therefore, this model has a spatial structure suitable for expressing a landscape in a spherical image.

References

[1] Barazzetti, L., Oteri, A. M., Pellicanò, R., Santoro, A., & Valiante, C. RENDERING EQUIRECTANGULAR PROJECTIONS ACQUIRED WITH LOW-COST 360° CAMERAS. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **XLVI-2/W1-2022** (2022). 57-63

[2] Cai, Q., Li, M., Ren, D., Lyu, J., Zheng, H., Dong, J., & Yang, Y. H. Spherical pseudo-cylindrical representation for omnidirectional image super-resolution. In Proceedings of the AAAI Conference on Artificial Intelligence **38**, 2 (2024). 873-881

[3] Sui, X., Wang, S., & Fang, Y. A Survey on Objective Quality Assessment of Omnidirectional Images. In Proceedings of 2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) IEEE (2024) 1-6

[4] D'Amato, R., Cutolo, F., Badiali, G., Carbone, M., Lu, H., Hogenbirk, H. & Ferrari, V. Key Ergonomics Requirements and Possible Mechanical Solutions for Augmented Reality Head-Mounted Displays in Surgery Multimodal Technol. Interact. **6(2)** (2022) 15

[5] Stefana, E., Marciano, F., Rossi, D., Cocca, P. & Tomasoni, G., Wearable Devices for Ergonomics: A Systematic Literature Review, Sensors **21(3)** (2021) 1–24

[6] Lanyi, C. S. & Withers, J. D. A. Striving for a Safer and More Ergonomic Workplace: Acceptability and Human Factors Related to the Adoption of AR/VR Glasses in Industry 4.0 Smart Cities 2020 **3** (2020) 289-307

[7] Lang, S., Dastagir Kota, M. S. S., Weigert, D. & Behrendt, F., Mixed reality in production and logistics: Discussing the application potentials of Microsoft HoloLensTM, Procedia Comput. Sci. **149** (2019) 118–129

[8] Burmaoglu, S., Trajkovik, V., Tutukalo, T. L., Yalcin, H. & Caulfield, B., Evolution Map of Wearable Technology Patents for Healthcare Field, Wearable Technol Med. Heal. Care (2018) 275–290

[9] Ando, T., Lee, J., Shinozaki, M., Kitajima, T., An, Q., & Yamashita, A. Efficient Distortion Mitigation in Equirectangular Images for Two-View Pose Estimation. International Journal of Automation Technology **19(3)** (2025) 226-236

[10] O. -J. Kwon, J. Cho & S. Choi A Center-to-Edge Progression for Equirectangular Projected 360° JPEG Images. IEEE Access **9** (2021) 6921-6929

[11] Yang, W., Qian, Y., Kämäräinen, J. K., Cricri, F., & Fan, L. Object detection in equirectangular panorama. In Proceedings of 24th international conference on pattern recognition (icpr) IEEE (2018) 2190-2195

[12] McNamee, J., Debattista, K., & Chalmers, A. Efficient Remote Rendering Using Equirectangular Projection CGVC (2017) 53-59

[13] Erkelens, C.J. Geometric Constraints of Visual Space. i-perception. **12(6)**:1-12 2021

[14] Walch, M., Frommel, J., Rogers, K., Schüssel, F., Hock, P., Dobbelstein, D., & Weber, M. Evaluating VR driving simulation from a player experience perspective. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (2017) 2982-2989

[15] Araújo, A. B. Workshop-drawing equirectangular perspectives for VR panoramas with Eq A sketch 360. In Proceedings of 2021 7th International Conference of the Immersive Learning Research Network (iLRN) IEEE (2021) 1-3

[16] Nishimoto, H. Geometric model in visual space. JSIAM Letters. **15** (2023). 105-108

[17] Chen, Y., Wang, X., & Xu, H. Human factors/ergonomics evaluation for virtual reality headsets: a review. CCF Transactions on Pervasive Computing and Interaction **3(2)** (2021) 99-111

[18] S. Ko & L. H. Finkel. Spatial-frequency analysis in the perception of perspective depth. Netw. Comp. Neural Syst. **8**, 3 (1997) 335–352

[19] C. Ronchi, R. Iacono & P.S. Paolucci, The "Cubed Sphere": A new method for the solution of partial differential equations in spherical geometry. J. Comput. Phys., **124**, 1 (1996) 93-114.

[20] Zhou, G., Li, H., Song, R., Wang, Q., Xu, J., & Song, B. Orthorectification of fisheye image under equidistant projection model. Remote Sensing, **14**, 17 (2022) 4175.