Learning 4-D Spatial Representations Through Perceptual Experience With Hypercubes

Takanobu Miwa, Yukihito Sakai, and Shuji Hashimoto, *Member, IEEE*

***Abstract*—Imagine a day when humans can form mental representations of higher-dimensional space and objects. These higher-dimensional spatial representations may enable us to gain unique insights into scientific and cultural advancements. To augment human spatial cognition from three to four dimensions, we have developed an interactive 4-D visualization system for acquiring an understanding of 4-D space and objects. In this paper, we examine whether humans are capable of formulating 4-D spatial representations through perceptual experience in 4-D space with 4-D objects. Participants learn about 4-D space and hypercubes through an interactive system, and are then examined on a series of 4-D spatial ability tests. They demonstrate the ability to perform perspective taking, navigation, and mental spatial transformation tasks in 4-D space. The results provide empirical evidence that humans are capable of learning 4-D spatial representations. Moreover, the results support the interpretation that humans form a cognitive coordinate system, consisting of an origin and four directional axes, to understand 4-D space and objects.**

***Index Terms*—4-D interaction, perspective taking, spatial representations, spatial transformations, spatial visualization ability.**

I. Introduction

MENTAL representations of space and objects are strongly related to human cognitive processing, which includes thinking, learning, and problem solving. For instance, the ability to understand shapes, sizes, orientation, and spatial relations, to use mental maps to orient oneself in a maze-like environment, and to imagine different perspectives of an object is rooted in these mental representations. Furthermore, spatial representations are important in nearly all areas of science, technology, engineering, and mathematics, because spatial thinking is widely used as a tool for learning and development. For example, students of these disciplines might draw figures and diagrams to grasp the relationships among variables when solving a problem or proving a theorem. Thus, for many psychologists, cognitive scientists, and educators, it is of both theoretical and practical importance to study the cognitive structure and development of human spatial representations.

As we live in a world with three dimensions and interact with 3-D objects, we have evolved a vision system and spatial cognition that is adapted to 3-D space and objects. For many researchers, this naturally makes mental representations of 3-D space and objects the primary area of interest [1]–[8]. However, it is reasonable to consider whether the dimensional limitations of the physical world necessarily restrict human spatial representations to three dimensions. In other words, the question of whether humans can acquire mental representations of higher-dimensional space and objects without relying on mathematical representations remains an open question. As science and technology cannot make advances without an understanding of complex higher-dimensional data, and because higher-dimensional spatial representations enable us to gain direct insights into such data, this fundamental question merits greater attention.

To answer this question, one possible approach comes from the theory of empiricism. Empiricist philosophers consider all human knowledge to be primarily derived from sensory experience. According to this view, our mental representations of 3-D space are formed from perceptual experiences of and interactions with 3-D objects. Assuming this is true, if we can accumulate experience of higher-dimensional space in an environment that provides interactions with higher-dimensional objects, we will acquire mental representations of higher-dimensional space and objects. Indeed, some studies have demonstrated the validity of this hypothesis experimentally [9]–[14]. These studies examine human 4-D spatial ability by measuring judgments of simple 4-D tasks and provide important evidence that humans are capable of acquiring 4-D spatial representations. Hence, our current research interest continues to more complex 4-D judgments, such as spatial orientation, perspective taking, and spatial transformations.

In this paper, building on our preliminary work [15], we examine the possibility of acquiring 4-D spatial representations by testing participants' abilities to perform perspective taking and mental spatial representations in 4-D space. In the experiments, we provide participants with extensive training of 4-D space and objects through the interactive 4-D visualization system developed in our previous work [16]–[18]. We then use two different spatial recognition tasks to measure the participants' 4-D spatial visualization ability. In the first task, which

Manuscript received June 7, 2016; revised January 14, 2017 and April 20, 2017; accepted May 24, 2017. Date of publication June 1, 2017; date of current version June 8, 2018. This work was supported in part by Waseda University Grant for Special Research Projects (Project number: 2015S-084), and in part by the Waseda University Future Robotics Organization. *(Corresponding author: Takanobu Miwa.)*

T. Miwa and S. Hashimoto are with the Faculty of Science and Engineering, Waseda University, Tokyo 169-8555, Japan (e-mail: takmiwa@shalab.phys.waseda.ac.jp).

Y. Sakai is with the Faculty of Engineering, Fukuoka University, Fukuoka 814-0180, Japan.

This paper has supplementary downloadable multimedia material available at http://ieeexplore.ieee.org provided by the authors.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCDS.2017.2710420

assesses perspective taking and navigational skills, the participants guide themselves in 4-D space to obtain the desired view of a 4-D object. In the second task, which assesses mental spatial transformations, we ask the participants to mentally reconstruct a 4-D object from multiple 3-D projection drawings. The abilities required for these tasks involve mental representations of the spatial relations between an individual and objects in 4-D space. Thus, successful performance by the participants in these 4-D tasks would provide empirical evidence that humans are able to learn 4-D spatial representations through perceptual experiences.

The remainder of this paper is organized as follows. First, we review the relevant literature on 4-D spatial cognition and state the position of this paper in Section II. In Section III, we present the configuration of the interactive visualization system used in the evaluation experiments. In Section IV, we describe the methods and results of the first experiment, which assesses perspective taking and navigational skills in 4-D space. We describe the method and results of the second evaluation experiment, which assesses the ability to perform mental spatial transformations in 4-D space, in Section V. A general discussion of the results and their implications is given in Section VI. Finally, we conclude this paper in Section VII.

II. Related Work

In this section, we review the relevant literature on 4-D spatial cognition. We first summarize methods from the fields of computer graphics and virtual reality for visualizing and interacting with 4-D objects, and then discuss some recent studies on human 4-D spatial cognition from the field of psychology.

A. Visualization of and Interaction With 4-D Space and Objects

Conventional 4-D visualization techniques take one of two major approaches. The first projects a 4-D object into 3-D space, just as 2-D projections of 3-D objects are formed on the retina [19]–[24]. The second approach slices a 4-D object with a hyperplane in 4-D space, just as we cut a 3-D object with a 2-D plane [24]–[26]. In this paper, we focus on the first approach. This has the advantage that it maintains various original 4-D geometric features, not only the structural continuity and spatial relations, but also parallelism and orthogonality. This advantage helps people determine the overall shape of a 4-D object and geometric features, such as its size, position, and orientation. In previous studies, a 4-D eye-point is either fixed or undergoes limited changes. Here, we construct a visualization algorithm that enables 4-D objects to be observed from an arbitrary position and direction in 4-D space.

There have been some efforts to develop 4-D interaction techniques. In these systems, the operation of common input devices, such as a mouse, keyboard, or touch-screen [27], [28] was associated with geometric operations on a 4-D object. Another approach associates human body motion with geometric operations on 4-D objects [29]. These systems enable the user to observe the rotation of a 4-D object in 3-D space. However, because the interface in these studies was designed in 3-D space, the association between the user's operation in 3-D space and the geometric operation in 4-D space is determined by art rather than nature. Therefore, we have developed a new interface that enables the user to control a 4-D eye-point and a 4-D viewing direction in a natural fashion using the geometric processing of 4-D space.

In our previous work, we constructed a 4-D visualization algorithm via 5-D homogeneous processing. This algorithm enables 3-D perspective drawings of any 4-D data to be visualized in 3-D space from an arbitrary 4-D eye-point, viewing direction, and viewing field [30], [31]. Moreover, we constructed an interactive 4-D space display system that translates human actions in 3-D space to the movement and direction of a 4-D eye-point bounded on a 4-D spherical surface [30]. We associated human actions in 3-D space with an intuitive interface, in this case, a flight-controller pad. With this system, users can observe 4-D data, such as 4-D solids, 3-D time-series data, and 4-D mathematical data in an arbitrary 4-D viewing field while intuitively moving in 4-D space [30], [32], [33]. In addition, we generalized 4-D geometric element definitions and interference via 5-D homogeneous processing [34].

Through these studies, we have found that people utilize principal vanishing points as landmarks to understand their position and orientation as they move in 4-D space. Inspired by this discovery, we developed a novel algorithm that determines the 4-D eye-point and 4-D viewing direction from the 3-D positions of the principal vanishing points. Using this algorithm, we constructed a new interactive 4-D visualization system that employs the principal vanishing points as an interface to control the movement of a 4-D eye-point and changes in the 4-D viewing direction [16]–[18]. To evaluate the effectiveness of our interactive system based on principal vanishing points, we compared the usability of our system with that of two conventional interaction methods: 1) a classic keyboard-based system, which handles parameter changes regarding the 4-D eye-point movement and 2) our previous system, which utilizes a flight-controller pad associated with human actions in 3-D space. We performed an objective evaluation of the operation time and subjective evaluation of how it feels to operate the system. The results suggest that our interactive system has superior usability in terms of 4-D eye-point control to observe various 3-D perspective drawings of 4-D data [16]. Therefore, in this paper, we use this interactive system as an apparatus to train experimental participants and examine their 4-D spatial ability.

B. 4-D Spatial Cognition

As discussed above, research has shown that 4-D space and objects can be visualized in 3-D space with the aid of computer graphics and virtual reality. However, it remains to be determined whether humans can acquire mental representations or an intuitive understanding of 4-D space and objects without the aid of mathematical representations.

There are some studies that challenge this possibility. They can be classified into two categories according to how they assess 4-D spatial representations. The first category relies on informal subjective reports that assess the participants' ability to acquire an understanding of 4-D space and objects. For example, Davis *et al.* [9] reported that mathematicians who interacted with a hypercube visualized by a computer claimed that they suddenly "felt" it. These subjective reports have significant importance as initial evidence of the capability to develop mental representations of 4-D space and objects. Nevertheless, informal subjective reports cannot quantitatively reveal what type of 4-D spatial representations people acquire from the 4-D experience. Thus, studies including objective evaluations are needed to probe the possibility and nature of 4-D spatial representations.

The second category relies on objective evaluation methods that assess performance on cognitive tasks related to 4-D space and objects. As it is assumed that all spatial tasks are solved using internal manipulations of mental images, the ability to solve a certain task within 4-D space can serve as evidence that a person has 4-D spatial representations in their mental space. Aflalo and Graziano [10] used path integration as a test of 4-D navigational skills, because successful path integration requires a mental map of an environment. In the study, participants used a keyboard to navigate from a starting point to an end point in a mazelike environment. They then indicated the direction from the end point back to the starting point by changing the 4-D viewing direction until they were facing the starting point. The results show that the participants learned to indicate the correct direction as they gained 4-D experience over multiple trials. Ambinder *et al.* [11] used spatial judgment tasks consisting of distance and angle estimation in 4-D space. In their study, participants examined 3-D slices of a 5-cell that were displayed in a CAVE-type virtual reality system. They then formed spatial judgments about the geometric features of the 5-cell, including the distance between two vertices and the angle between two edges. The results indicate a positive correlation between the participant responses and the correct distance and angle, which suggests that participants were able to make judgments about these 4-D properties. A follow-up study [12] in which the visualization method was switched from a slicing approach to orthogonal projection confirmed this finding.

Although people have demonstrated the ability to form certain judgments about 4-D objects, this does not in itself guarantee that their mental representation is actually grounded in four dimensions. People may use a variety of strategies, such as prior intuition, a trial-and-error approach, and mechanical solutions, which are useful for solving the task but not based on 4-D spatial cues. This is always a problem when assessing 4-D spatial ability, and it even often occurs for 3-D imagery [35]–[37]. For this reason, when we examine 4-D spatial representations through objective evaluation, it is necessary to design a 4-D spatial ability test such that it can never be solved by such strategies. If it is difficult to design such a task, the solver's strategy can be identified as a 4-D one by behavioral observation, analysis of the experimental results, and post-experiment questionnaires and interviews.

Wang [13] identified this problem and defined the key conditions that 4-D spatial representations should satisfy. According to his assertion, 4-D spatial representations can be defined as perceptual or cognitive representations of 4-D objects or environments that can support judgments about 4-D spatial relations or spatial properties without using definition-based lower-dimensional solutions, algebraic equations, or feedback training. Using these criteria, Wang [14] examined 4-D spatial representations using the hypervolume, which is a geometric property unique to 4-D space, as a test of 4-D object visualization ability. In the study, the task solvers observed the 3-D orthogonal projection of a randomly shaped 5-cell that horizontally rotated to a depth direction in 4-D space. They then provided their answer for the hypervolume of the 5-cell by adjusting the size of a hyperblock to match that of the 5-cell. The results show a positive correlation between the solvers' responses and the correct hypervolume, but not the definition-based, lower-dimensional cues, and Wang concluded that the participants' 4-D spatial representations meet the abovementioned definition of 4-D spatial representations. This suggests that humans are able to form some sort of 4-D spatial representations that help them perform object visualization.

Here, we briefly summarize the common points and differences between previous research and this paper. Considering the characteristics of these previous studies, in this paper, we examine human 4-D spatial representations under the following conditions.

1) *4-D Learning Method:* As Aflalo and Graziano [10] used extended training, we allow participants to practice with 4-D space and objects until they are satisfied with their own understanding. Moreover, unlike Wang [13], who restricted 4-D rotation to the depth direction, we allow participants to observe the 4-D objects from arbitrary 4-D positions and directions. These conditions help the participants to acquire their own 4-D spatial representations and make more complex 4-D judgments than previously shown.
2) *Experimental Tasks:* Ambinder *et al.* [11] and Wang [14] examined participants' understanding of the geometric properties of an object. In contrast, we examine 4-D spatial visualization ability, which corresponds to the ability to process spatial relations, movements, and different perspectives. We therefore use spatial tasks that can be solved by perspective taking, navigational skills, and mental spatial transformations, rather than simple tasks, such as distance, angle, and hypervolume estimation. This enables us to verify whether 4-D spatial representations are sufficiently flexible and uniform to perform active cognitive processing, such as predictions of visual changes in 4-D objects, behavioral decisions in 4-D space, and mental creation of a novel 4-D object. Moreover, this helps us develop hypotheses about possible forms of 4-D spatial representations.
3) *Data Analysis:* As in conventional studies, we evaluate participants' performance based on individual subject analysis to confirm whether they have solved the experimental tasks with 4-D strategies or mechanical strategies. We evaluate each participant's ability by observing their actions in 4-D space and scoring their performance on the experimental tasks. This enables us to verify whether the participants' mental representations are actually grounded in four dimensions.

III. Interactive 4-D Visualization System

In this section, we explain the interactive system constructed in our previous work [16]–[18]. All experiments reported in this paper were conducted using this system.

A. Basic Concept of the Interactive System

To understand the shape of a 3-D object, humans move around the object to observe it from various directions in 3-D space. Similarly, to understand the surrounding environment, we look around and move in the environment. By analogy, we consider that, if we can move in 4-D space and observe 4-D objects from various positions, directions, and distances, we should be able to acquire an intuitive understanding of 4-D space and objects. Based on this idea, in the interactive system, we display various 3-D perspective drawings of 4-D objects from an arbitrary 4-D eye-point and viewing direction. The 4-D eye-point can be interactively controlled to move along a 4-D spherical surface that surrounds the 4-D object. For this interaction, we focused on using the principal vanishing points as an interface for 4-D eye-point control.

Before considering the 4-D case, we begin with a study of the 3-D case. As stated by Foley *et al.* [38], in 3-D space, a perspective projection is a method of mapping 3-D points to a projection plane. The projection points are obtained as the intersections of straight projection rays with the projection plane. The straight projection rays, called projectors, are formed by connecting the center of projection, called the eye-point, and each of the 3-D points. The projection plane, called a 2-D screen, floats in front of the eye-point and is perpendicular to the viewing direction. The perspective projection shows distant objects as being smaller than near objects. This is similar to the human eye and camera lenses; therefore, the perspective projection produces a realistic representation of a 3-D object on the 2-D screen. To be more precise, when any parallel lines in 3-D space are not parallel to the 2-D screen, their perspective projections will converge toward a single "vanishing point" on the 2-D screen. In 3-D projective space, the parallel lines intersect at infinity. Hence, the vanishing point is the projection of the point at infinity associated with the parallel lines. If the parallel lines are parallel to one of the three principal coordinate axes, the vanishing point is specifically referred to as a "principal vanishing point." In the 3-D case, one, two, or three principal vanishing points will appear on the 2-D screen, corresponding to the number of principal coordinate axes that are not parallel to the 2-D screen.

Each principal vanishing point is determined by the geometric relationship among the eye-point, viewing direction, and principal coordinate system. Conversely, it is possible to estimate the eye-point and viewing direction in 3-D space from each position of the principal vanishing points on the 2-D screen [39]. This implies that it is possible to control the eye-point in 3-D space by manipulating the position of the principal vanishing points on the 2-D screen, as shown in Fig. 1.

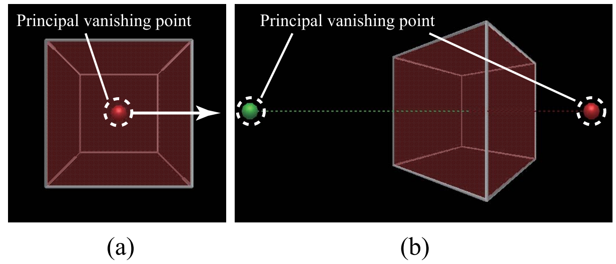


Fig. 1. Observation of a cube in 3-D space, operating on the principal vanishing points. (a) 2-D perspective drawing of the cube and the principal vanishing point. (b) Principal vanishing point after eye-point movement in 3-D space.

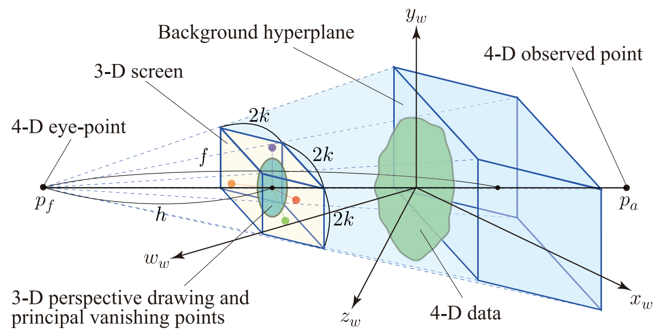


Fig. 2. Visualization model of 4-D space and objects.

B. 4-D Eye-Point Control by Operating on Principal Vanishing Points

We now explain interactive 4-D eye-point control by operating on the principal vanishing points. We extend the setup explained above from 3-D space to 4-D space. That is, we consider controlling the 4-D eye-point to move along a 4-D spherical surface, centered on the origin of the 4-D world-coordinate system, by handling the principal vanishing points displayed in 3-D space.

Fig. 2 shows the 4-D visualization model. A 3-D perspective drawing of a 4-D object is obtained by converting data defined in the 4-D world-coordinate system $x_w y_w z_w w_w$ to data in the 3-D screen-coordinate system $x_s y_s z_s w_s$ [30], [31]. The 4-D viewing direction is defined as the direction from the 4-D eye-point $p_f(x_{p_f}, y_{p_f}, z_{p_f}, w_{p_f})$ to the 4-D observed point $p_a(x_{p_a}, y_{p_a}, z_{p_a}, w_{p_a})$ in the 4-D world-coordinate system. The center of the 3-D screen and that of the background hyperplane are located at distances h and $f(> h)$, respectively, from the 4-D eye-point on the 4-D visual axis. The dimension of the 3-D screen is $2k \times 2k \times 2k$ in the 3-D screen-coordinate system. In contrast to visualization models in conventional studies, only 4-D objects that are inside the 4-D viewing field (defined as a truncated pyramid formed by the 4-D eye-point, 3-D screen, and background hyperplane) are visualized on the 3-D screen. The visualization algorithm includes a view field transformation, perspective transformation, and clipping operation in 4-D space using 5-D homogeneous processing. This framework can visualize any 4-D data, including points at infinity, onto 3-D space from an arbitrary 4-D eye-point and viewing direction.

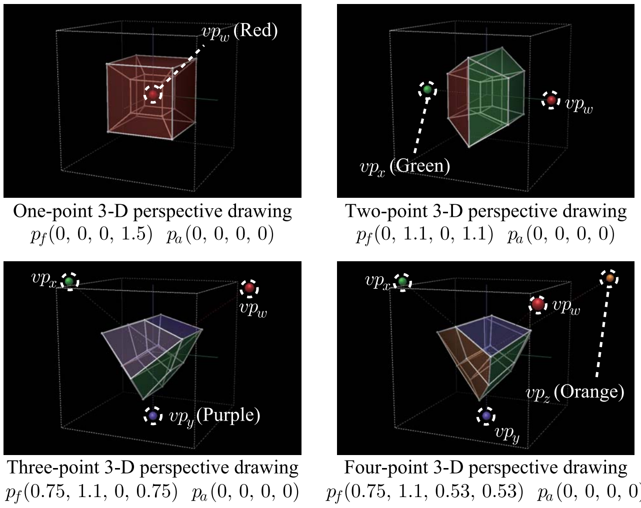


Fig. 3. Principal vanishing points and 3-D perspective drawing of a hypercube. The inside region surrounded by the white dashed wire-frame cube corresponds to the 3-D screen. [The parameters (k, h, f) for the 4-D viewing field are $(0.5, 0.5, 100)$.]

The principal vanishing points vp_x, vp_y, vp_z, and vp_w are represented by projecting points at infinity in the x_w-, y_w-, z_w-, and w_w-directions, respectively, onto 3-D space. As shown in Fig. 3, the number of principal vanishing points and their 3-D positions are determined by the spatial relationships among the 4-D eye-point, 4-D viewing direction, and 4-D world-coordinate system. Accordingly, it is assumed that there is a converse relation such that the 4-D eye-point can be derived from the positions of the principal vanishing points. Based on this assumption, we constructed an interaction algorithm that makes position changes of the principal vanishing points in 3-D space correspond to 4-D eye-point movement in 4-D space [16]. This interaction algorithm is composed of two processing steps. As the user picks and moves one principal vanishing point in 3-D space, the first processing step estimates the other principal vanishing points from the principal vanishing point being operated on. The second processing step estimates the parameters of the 4-D viewing direction from the principal vanishing points to determine the 4-D eye-point. With this interaction algorithm, the user can intuitively control the 4-D eye-point by manipulating the position of the principal vanishing points in 3-D space while using them as landmarks to recognize his/her viewing position in 4-D space (for more details of the algorithms, see Appendix A).

Let us add a complementary explanation of the implementation of these algorithms. In the 4-D visualization algorithm, the 4-D view field transformation enables movement of the 4-D eye-point and changes in the 4-D viewing direction from the 4-D eye-point to the 4-D observed point [30], [31]. Thus, when we integrate the interaction algorithm into the visualization algorithm, we initially introduce the 4-D eye-point determined from the second processing step of the interaction algorithm and the 4-D observed point into the 4-D view field transformation. We then recalculate the movement of the 4-D eye-point and the change in the 4-D viewing direction from the 4-D eye-point to the 4-D observed point. However, part of this algorithmic procedure is modified when we implement the algorithms in the interactive system. In this case, we do not apply the 4-D eye-point and the 4-D observed point, but rather the 4-D eye-point and the parameters of the 4-D viewing direction derived from the second processing step of the interaction algorithm to the 4-D view field transformation. Although this modified 4-D view field transformation may differ from the 4-D view field transformation of the original algorithm depending on the movement history of the 4-D eye-point, in our implementation, this modification has a positive effect on the style of the 4-D observation. If we implement the visualization algorithm without any modification, the attitude of the 4-D eye-point is restricted, as the 4-D upper direction of the 4-D eye-point is always in the upward vertical direction in 4-D space. This restriction causes the attitude of the 4-D eye-point to change at some positions on the 4-D spherical surface, regardless of the user's will. (For a specific example, see Appendix B.) In this state, the continuity of 4-D observation is not ensured, and the observation becomes unnatural. To avoid this irregularity in 4-D observation, the movement history of the 4-D eye-point can be considered in deriving the parameters of the 4-D viewing direction from the principal vanishing points, and the attitude of the 4-D eye-point is maintained before and after its movement. As a result, our implementation enables the user to observe 4-D space and objects while he/she continuously moves the eye-point along the 4-D spherical surface in a natural style.

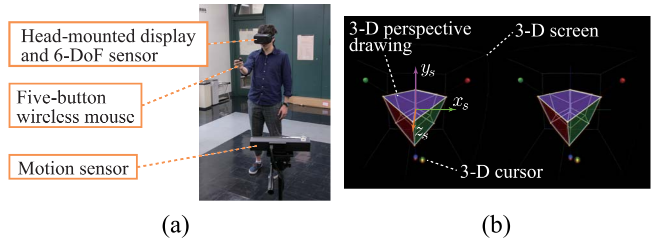


Fig. 4. Example of actual use of the interactive system. (a) Configuration of the system. (b) Example of the stereoscopic image seen through the head-mounted display.

C. Configuration of the Interactive System

Fig. 4 shows a configuration of the interactive system. The system consists of commercially available products: a personal computer, a head-mounted display with a built in 6-DoF sensor, a motion sensor, and a five-button wireless mouse. The 3-D virtual space seen through the head-mounted display coincides with the $x_s y_s z_s$-space in the 3-D screen-coordinate system $x_s y_s z_s w_s$. The movable region for two of the principal vanishing points, vp_x and vp_y, is restricted to the $x_s y_s$-plane and the y_s-axis, respectively, whereas the other principal vanishing points, vp_z and vp_w, can move freely in 3-D space. These restrictions mean that the 4-D upper direction of the 4-D eye-point is maintained in an upward or downward vertical direction in the 4-D world-coordinate system during the interaction. The 3-D virtual space has the same scale as real space, and the viewing position and viewing direction in 3-D virtual space are associated with the user's head position and orientation, respectively; likewise, the user can observe a 3-D perspective drawing from any direction. The 3-D perspective drawing, principal vanishing points, and 3-D screen are displayed in 3-D virtual space as a stereoscopic image. The 3-D screen has a size of 300 mm × 300 mm × 300 mm, and is placed 1.5 m off the ground according to the user's request. Under these conditions, the user observes various 3-D perspective drawings of the 4-D object while handling the principal vanishing points in 3-D space with simple pick-and-move operations using a 3-D cursor associated with their hand motion.

IV. Experiment 1: Perspective Taking

To examine whether humans can form 4-D spatial representations, we conducted two different experiments that assess the ability to perform spatial tasks in 4-D space. In this section, we describe the first experiment, which assesses perspective taking skills and navigational skills in 4-D space.

A. Methods

We first introduce the participants, apparatus, and stimuli of the experiment.

1) Participants: Eight male participants and four female participants were tested. Their mean age was 20.5 years. Seven participants were students or graduates in science and engineering courses, and the others were students in arts courses. None of the participants had knowledge or experience of 4-D space and objects, and had no prior knowledge of the design of the experimental tasks.

2) Apparatus: We used a personal computer (Intel Core i7 3.90 GHz, NVIDIA GeForce GTX 680, 32-GB RAM) installed with Windows 8.1 (Microsoft), a head-mounted display (Oculus Rift DK2), a motion sensor (Microsoft Kinect v2), and a wireless mouse (ELECOM M-GE3DL). The software was implemented in C#, OpenTK, and the SDKs of the system components under Visual Studio (Microsoft). The system guarantees real-time performance (75 frames/s) during the experimental tasks. The participants used the head-mounted display and the five-button wireless mouse to accomplish experimental tasks performed in the interactive system. Although the head-mounted display and motion sensor have a small latency, this did not affect the participants' performance.

3) Stimuli: We used a hypercube as the observation object. A hypercube is the 4-D analog of a cube, and consists of eight cubical cells. In total, a hypercube has 16 vertices, 32 edges, 24 surfaces, and 8 cells. Four of the eight cells are connected at each vertex, three of those are connected at each edge, and two of those are connected at each surface. Opposite cells are parallel to each other, and consecutive cells form a right angle with each other in 4-D space. In the experiment, the coordinates of the vertices of the hypercube are given as permutations of ($\pm0.5, \pm0.5, \pm0.5, \pm0.5$). We assigned a different color (red, pink, green, cyan, purple, white, orange, or yellow) to each cell of the hypercube and placed it at the origin of the 4-D world-coordinate system such that each of the eight cells was located in a different positive or negative position on each axis of the 4-D world-coordinate system, at a distance of 0.5 from the origin. To clearly visualize the edges inside the 3-D perspective drawings of the hypercube, the 3-D perspective drawings were rendered with a semitransparent surface and a reticular stipple pattern. During the experiment, the participants interacted with this hypercube using the interactive system.

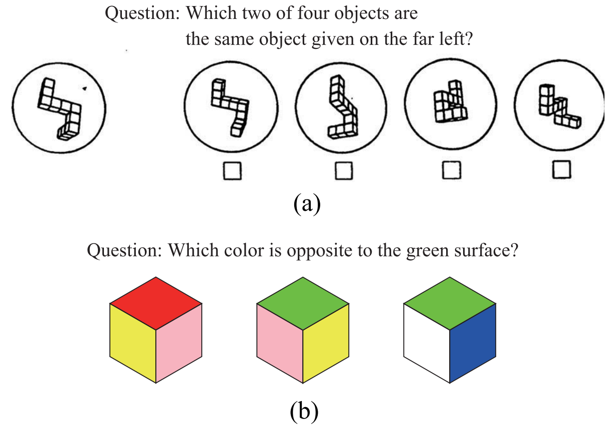


Fig. 5. Examples of the preliminary test. (a) MRT. The second and third options from the left are correct. (b) CCT. "Red" is the correct answer.

B. Procedure

Each participant took a preliminary test and filled out a self-profiling form at the beginning of the evaluation experiment. This took approximately 30 min. Participants then received a brief lecture on some basic aspects of 4-D space and objects, the 4-D visualization model, and how to use the interactive system based on an analogy with 3-D space. This explanation took approximately 15 min.

After the preliminary test and lecture, the participants underwent a 4-D learning period. We first gave the participants approximately 10 min to practice using our system and make personal adjustments. The participants then studied 4-D space and the hypercube to acquire 4-D spatial representations. The length of the learning period varied from 120 to 180 min for each participant. Including breaks, the learning period took approximately 140–210 min in total.

Finally, the participants proceeded to the task period, which took approximately 60 min. The participants were asked to handle the navigational tasks in 4-D space based on their 4-D spatial representations. To check the strategies used by participants for the experimental tasks, we interviewed them after they had finished the test.

Because the total duration of the experiment was too long to run in a single day, we limited the participants' daily operating time to 120 min in consideration of their tiredness. If participants were in the middle of the experimental work when this time limit was reached, they suspended the task and resumed it the next day.

1) Preliminary Tests: As individual differences in 3-D spatial perception could affect the development of 4-D spatial representations, we ran a preliminary test to determine the participants' intrinsic spatial ability. Fig. 5 shows examples

TABLE I
RESULTS OF THE PRELIMINARY TESTS

Survey items	A	B	C	D	E	F	G	H	I	J	K	L
	Participant											
Spatial ability												
MRT	2.4	5.6	3.6	3.9	2.0	2.4	3.1	1.9	3.0	2.7	2.7	1.9
CCT	1.7	2.7	2.2	2.4	0.7	1.5	2.2	1.7	1.9	1.2	2.1	1.3
Spatial Confidence												
SBSOD	2.5	4.4	2.8	2.7	1.9	1.3	1.5	2.3	1.8	2.7	3.1	3.3
VISQ	2.4	3.1	3.0	2.8	1.6	1.3	2.3	2.0	1.6	2.5	2.9	3.5
Profile												
Age	20	20	20	19	20	19	20	19	20	27	22	20
Sex	Male	Male	Male	Male	Male	Female	Female	Female	Female	Male	Male	Male
Academic background	Science	Art	Science	Science	Art	Art	Science	Science	Science	Science	Science	Art

of the preliminary test. We used 20 trials of the redrawn version of the mental rotation test (MRT) [40] and ten trials of the color cube test (CCT), which we developed ourselves. For both tests, we recorded the number of correct responses and the total response times. The overall score was calculated by dividing the number of correct answers by the total response time; thus, the higher the score, the better the spatial ability. In addition to these two spatial ability tests, we used self-reporting to determine the participants' spatial confidence. We used the Santa Barbara sense-of-direction scale (SBSOD) [41], which consists of 15 questions, and the visual imagery style questionnaire (VISQ) [42], which consists of 12 questions. For both tests, the participants were asked to respond to each question on a scale of 1–5. The overall score was calculated as the average over all responses, where a higher score indicates better spatial confidence.

Table I presents the results of the preliminary test. The participants exhibited a range of 3-D spatial ability and spatial confidence. We discuss whether these individual variations affected the results of the 4-D spatial ability tests in Section VI.

2) Learning Tasks: In the learning task, participants were asked to observe and understand 4-D space and the hypercube. Using the interactive system, each participant observed the hypercube from various positions in 4-D space while moving freely around the hypercube. The participants were allowed to work on the learning task until they were satisfied with their understanding, up to a time limit of 180 min. The participants were allowed a 5-min break every 30 min. An illustration of the learning task can be found in the video clips in the supplementary material.

3) Perspective Taking and Navigational Tasks: In the test, participants were asked to guide themselves to a 4-D checkpoint position. The test consisted of two stages. Each stage had one practice trial followed by ten experimental trials. In the first stage, participants started from a 4-D position on the w_w-axis and moved toward one checkpoint. Thus, the participants visited ten checkpoints in total. The checkpoints were randomly selected from the 4-D positions from which three or four cells of the hypercube could be seen at once. In the second stage, participants again started from a 4-D position on the w_w-axis, and this time traveled through four checkpoints in order. Thus, the participants visited 40 checkpoints in total. The checkpoints were randomly selected from the 4-D positions from which one, two, three, or four cells of the hypercube could be seen at once. For each trial, the target checkpoint was described orally as a set of cell colors that participants would see at the checkpoint, for example, "go to a position in 4-D space from which you can see red, green, purple, and orange cells simultaneously." To solve this task, participants needed to imagine a point from which they could obtain the desired perspective of the hypercube and identify a reasonable route to the checkpoint in 4-D space, based on their mental representations of 4-D space and the hypercube. If the participants guided themselves to the checkpoints smoothly without losing spatial orientation in 4-D space, they were considered to have successfully acquired 4-D spatial representations.

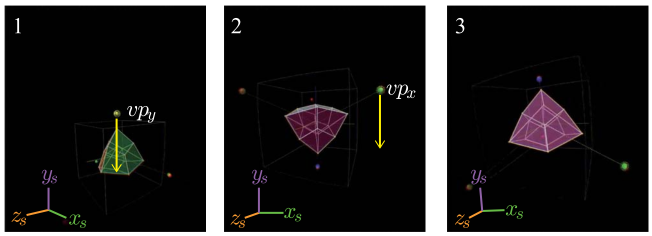
C. Results

This experiment was designed around an individual subject analysis based on behavioral observations. The experimental score is the number of target checkpoints that the participants moved to from the starting point or the previous checkpoint without losing track of their position and the direction of the target checkpoint in 4-D space. Considering the tendencies exhibited by the participants in our preliminary research [15], we defined criteria to determine whether the participants had become lost in 4-D space. That is, the participants should not arrive at the checkpoint by a random walk approach, they should not travel back and forth to the same position many times, and they should not move in the wrong direction once they were fairly close to the target checkpoint. For each trial, we carefully observed the participants' operations on the principal vanishing points in 3-D space and the moving path of the 4-D eye-point in 4-D space, and judged whether the navigation was a success or failure.

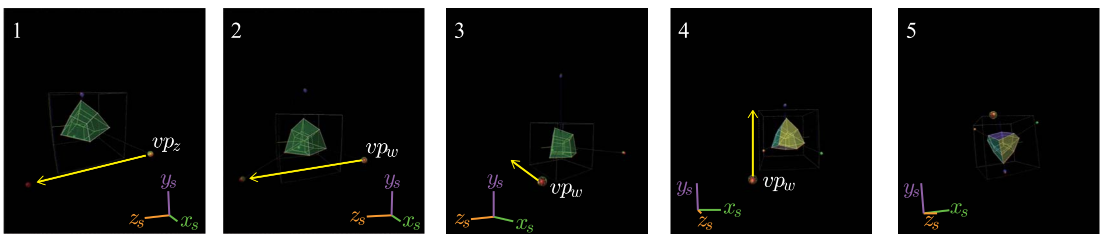
Table II presents a summary of the learning time, score, and participants' strategy toward the experimental tasks. Overall, eight participants (A, B, C, D, G, J, K, and L) navigated smoothly in both stages, which suggests that they were able to perform perspective taking and acquire navigational skills in 4-D space. Three participants (E, H, and I) exhibited worse performance in the first stage or both stages. Participant F dropped out of the experiment after the learning task, because she was not able to form a mental image of 4-D space and the hypercube.

TABLE II
RESULTS OF EXPERIMENT 1

Performance	Participant A	B	C	D	E	F	G	H	I	J	K	L
Learning time [min]	130	180	180	120	180	180	150	180	180	180	180	160
Score												
First stage	9 / 10	8 / 10	7 / 10	7 / 10	5 / 10	—	8 / 10	2 / 10	2 / 10	7 / 10	9 / 10	8 / 10
Second stage	38 / 40	31 / 40	34 / 40	38 / 40	26 / 40	—	38 / 40	27 / 40	8 / 40	39 / 40	39 / 40	33 / 40
Strategy												
Direct flight	○	×	○	○	×	—	○	×	×	○	×	○
Relay-points	○	○	○	○	○	—	○	○	○	○	○	○
Trial and error	×	×	×	×	○	—	×	○	○	×	×	×



(a)



(b)

Fig. 6. Difference in principal vanishing point operations. (a) Direct flight strategy (participant D). (b) Relay-points strategy (participant A).

From our observations of the participants' actions in 3-D and 4-D space, we identified three strategies used in the experiment: the direct flight strategy, the relay-points strategy, and the trial-and-error strategy. Successful participants used the direct flight strategy and the relay-points strategy. The former involves moving directly along the shortest path to the checkpoint, whereas the latter involves creating paths by moving from view to view in a way that will surely lead to the target checkpoint. The participants who used these strategies were able to recover their spatial orientation even if they moved in the wrong direction or made a detour in 4-D space.

A significant difference between the direct flight strategy and the relay-points strategy is the number of operations on the principal vanishing points. Participants who guided themselves in 4-D space with the direct flight strategy performed fewer operations. In most cases, participants needed only one or two steps to accomplish the task. Another distinctive difference is the shape of the 4-D eye-point movement path. When we visualized the trajectory of the 4-D eye-point movement onto the 3-D screen with our interactive system, the trajectory of the direct flight strategy was shorter and had fewer bends than the other strategies. Considering these characteristics, we classified each successful trial manually based on behavioral observations of the participants' operations in 3-D space and visual observations of the trajectory of the 4-D eye-point movement.

We show a typical sample of the operations in Fig. 6; other results display similar trends. We use trial #26 of the second stage as an example. In this trial, participants moved from the starting point near $(0.75, -0.75, 0.75, 0.75)$ to the checkpoint near $(-0.75, 0.75, -0.75, -0.75)$. In other words, participants moved from the starting point to the other side of the 4-D spherical surface. Fig. 6 depicts screen shots of a series of principal vanishing point operations. Participant D used the direct flight strategy and accomplished the task in two steps, which was the minimum number of operations required. He simply moved the principal vanishing points vp_y and vp_x parallel to the y_s-axis. In contrast, participant A required four steps to reach the checkpoint. He moved the principal vanishing points vp_z and vp_w parallel to the z_s-, x_s-, and y_s-axis step-by-step using the relay-points strategy. In this strategy, participants create paths by the following easy-to-understand views in a way that is sure to reach the target checkpoint.

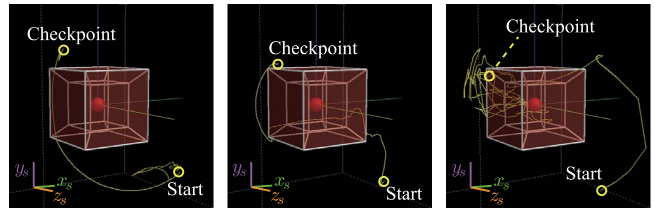
Fig. 7 shows the trajectories of the 4-D eye-point movements corresponding to the trial shown in Fig. 6. Using our interactive system, we draw the trajectory in 4-D space and visualize it on the 3-D screen from the 4-D eye-point $(0, 0, 0, 2.1)$ to the 4-D observed point $(0, 0, 0, 0)$. The parameters (k, h, f) of the 4-D viewing field are $(0.5, 0.5, 100)$. Note that we show the left view of the stereoscopic image displayed on the head-mounted display. Participant D's trajectory generally forms a unique smooth curve that represents the shortest path from the starting point to the checkpoint. Participant D passed through the zenith of the 4-D spherical surface as if it were an aerial passage following the polar route. In contrast, participant A's trajectory includes three bends. Each bend corresponds to the point at which the participant changed an operation on a vanishing point or changed the direction of movement of the principal vanishing points. Thus, participant A traveled through some relay points where he could confirm one or two cells included in the target perspective of the hypercube to ensure that he could reach the checkpoint.

According to the post-experiment interviews, all of the successful participants attempted to predict the direction of the target checkpoint. They also reported trying to plan a secure route from the starting point to the checkpoint by predicting the visual changes in the 3-D perspective drawing of the hypercube. For example, participants A and D reported that they used the direct flight strategy when they were required to move to the other side of the 4-D spherical surface and used the relay-points strategy when they needed to successively operate multiple principal vanishing points to reach a checkpoint. The successful participants had confidence in their strategy and understanding. Therefore, we are probably justified in thinking that the participants developed some sort of mental representations of 4-D space and the hypercube, and performed perspective taking and navigation in 4-D space.

In contrast, the unsuccessful participants relied heavily on the trial-and-error strategy. Fig. 7(c) visualizes participant E's trajectory. The zig-zag lines indicate that the participant got lost in 4-D space and moved the principal vanishing points at random in 3-D space. In particular, participants E and I exhibited little skill in operating principal vanishing points vp_z and vp_w in the z_s-direction, and they could not understand the 4-D eye-point movement in the 4-D direction corresponding to that operation. When participants observe a 3-D object from various directions in 3-D space, they need only move their eye-point vertically and horizontally around the object. Thus, the participants might naturally understand the correspondence between the 4-D eye-point movement and the principal vanishing point operations in the x_s- and y_s-directions. However, when the participants attempt to move in 4-D space, this understanding is insufficient. These participants probably saw the 3-D perspective drawing as a 3-D object that changed its appearance according to special transformation rules. As a result, they failed to understand the relationship between the principal vanishing point operations in 3-D space and the eye-point movement in 4-D space. We conclude that these participants failed to acquire 4-D spatial representations and that their understanding was insufficient for perspective taking and navigation in 4-D space.

V. Experiment 2: Mental Spatial Transformation

We received a positive insight into the possibility of participants using 4-D spatial representations from the results of Experiment 1. However, because Experiment 1 included some overlap between the learning tasks and the experimental tasks, it remained unclear whether the participants had acquired actual 4-D spatial representations or simply gained route knowledge of operations on the principal vanishing points. Thus, it was necessary to examine whether the participants' representations were universal and flexible enough for a novel 4-D spatial task. In this section, we describe a second experiment that tested the participants' ability to perform mental spatial transformations and manipulations of a 4-D object.



$p_f(0, 0, 0, 2.1)$ $p_a(0, 0, 0, 0)$ (a) | $p_f(0, 0, 0, 2.1)$ $p_a(0, 0, 0, 0)$ (b) | $p_f(0, 0, 0, 2.1)$ $p_a(0, 0, 0, 0)$ (c)

Fig. 7. Difference in the trajectory of the 4-D eye-point movement. (a) Direct flight strategy (participant D). (b) Relay-points strategy (participant A). (c) Trial-and-error strategy (participant E).

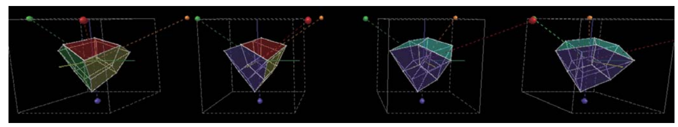


Fig. 8. Example of the stimuli.

A. Method

We first introduce the participants, apparatus, and stimuli of the experiment.

1) Participants: The 12 participants involved in Experiment 1 were invited to join Experiment 2.

2) Apparatus: We ran the experiment under almost the same conditions as Experiment 1, with the exception that, in this experiment, participants could not use the wireless mouse.

3) Stimuli: The stimuli were four four-point 3-D perspective drawings of one novel hypercube. Each stimulus was formed by randomly assigning a different color to each cell in the hypercube. The hypercube was then projected from four different 4-D viewing positions chosen from 16 candidates. As a result, these 3-D perspective drawings had the same exterior shape and different colorings. The four 3-D perspective drawings were arranged in a single line in 3-D virtual space. Using the interactive system, the participants observed the stimuli from various directions in 3-D virtual space to mentally reconstruct the entire image of the hypercube.

Fig. 8 shows an example of the stimuli (see video clip in the supplementary material). As in Experiment 1, to clearly visualize the inside of the 3-D perspective drawings, we rendered them with semitransparent surfaces with a reticular stipple pattern. Each of the 3-D screens was sized at 200 mm × 200 mm × 200 mm, and they were positioned 200 mm apart.

B. Procedure

After completing Experiment 1, the participants proceeded to Experiment 2. Each participant received an explanation of

TABLE III
RESULTS OF EXPERIMENT 2

Performance	Participant A	B	C	D	E	F
Scores						
Number of correct color-pairs	10 / 10	10 / 10	10 / 10	10 / 10	—	—
Number of correct answers	8 / 10	5 / 10	10 / 10	10 / 10	—	—
Confidence	5	3.6	4.6	4.3	—	—
Mean response time [s]	278 ± 91	564 ± 163	254 ± 248	325 ± 202	—	—
Solution						
Mental 4-D eye-point movement	○	×	○	○	—	—
Mental 4-D axes simulation	×	×	×	×	—	—

Performance	Participant G	H	I	J	K	L
Scores						
Number of correct color-pairs	9 / 10	10 / 10	4 / 10	10 / 10	10 / 10	8 / 10
Number of correct answers	9 / 10	5 / 10	4 / 10	10 / 10	7 / 10	8 / 10
Confidence	3.9	2.7	3.1	4.7	4.4	3.7
Mean response time [s]	279 ± 171	375 ± 65	264 ± 73	212 ± 84	578 ± 432	450 ± 279
Solution						
Mental 4-D eye-point movement	○	×	×	×	×	○
Mental 4-D axes simulation	×	×	×	○	○	×

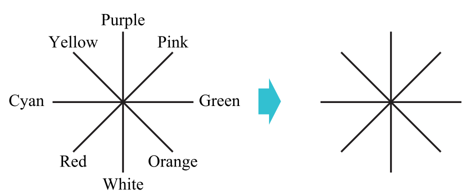


Fig. 9. Response sheet for the reconstructed coloring pattern.

the experimental task before completing one practice trial followed by ten experimental trials. To confirm the participants' strategies toward the task, they were interviewed after finishing all trials.

In each trial, the participants needed to mentally reconstruct a hypercube from the multiple 3-D perspective drawings displayed in 3-D virtual space. They were then asked to complete the response sheet shown in Fig. 9 to explain the coloring of the hypercube. Thus, this test is the 4-D version of the CCT shown in Fig. 5(b). Each participant allocated the eight colors to the graph on the right according to the analogy with the sample on the left, which represents the coloring pattern used in Experiment 1. Additionally, the participants rated their self-confidence in each response on a scale of 1–5, where 5 represents the positive end of the spectrum. This task requires the ability to transform and manipulate the 4-D structure of the hypercube from the given 3-D perspective drawings. Therefore, strong performance in this experimental task is an indicator of success in learning 4-D spatial representations.

This experiment was designed around an individual-subject analysis. The experimental score is the number of correct color-pairs, number of correct answers, and mean confidence rating for the ten trials. The number of correct color-pairs corresponds to the number of trials (out of ten) for which the participants correctly identified the four pairs of colors allocated to two cells facing each other in the hypercube. Note that we counted the answers as being correct even if they were rotated in 4-D space.

C. Results

The participants' performance and confidence ratings are presented in Table III. We summarize the score, mean response time, and participants' strategy. Overall, seven participants (A, C, D, G, J, K, and L) were able to mentally reconstruct the hypercube from its 3-D perspective drawings with a high degree of confidence. Three participants (B, H, and I) exhibited lower performance and confidence levels. Two participants (E and F) dropped out of the test because they could not imagine the structure of the hypercube at all.

As there are 210 possible coloring patterns, the probability of a participant successfully completing the task at random is $1/210$. The number of correct answers for participants other than E and F was higher than the random chance level. If the participants were able to find correct color-pairs in some way, they were finally required to decide whether the reconstructed hypercube was the normal object or a mirror object of the correct hypercube. In this case, the chance of correctly guessing at random is ultimately $1/2$. Six participants in the high performance group (A, C, D, G, J, and L) performed significantly better than the random chance level, suggesting that they were able to utilize their 4-D mental representations to solve the task. Although participant K's score only surpassed the level of random chance by a small margin, his performance was distinctive. He was wrong on the first three trials, and then gave the correct answer for seven consecutive trials. According to his comments, after the first three trials, he noticed that his answers were mirror objects due to a misunderstanding of the coloring pattern of the sample hypercube represented on the left side of the response sheet; he then corrected his mental representation. Thus, we consider that participant K should be included among the successful participants.

According to the post-experiment interviews, the successful participants seemed to solve the hypercube reconstruction test with a two-step approach. First, they found four sets of two cells that face each other in 4-D space. One solution is the elimination method; another is to focus on the principal vanishing points. If the participants understood that two such cells were on the same principal coordinate axis in 4-D space, they were able to determine the four sets by looking for the two cells corresponding to the same principal vanishing point.

The second critical step involves constructing the hypercube by appropriately arranging the four sets of two cells in 4-D space. One solution is to mentally simulate the 4-D eye-point movement to determine the 4-D spatial relationship among the four sets of two cells. For example, participant C reported that he simulated visual changes in the 3-D perspective drawing of the hypercube, starting from a one-point perspective drawing of each four-point 3-D perspective drawing, by mentally operating the principal vanishing points. Participants A, D, G, and L reported trying similar solutions. This approach is grounded in 4-D spatial thinking. In addition, the participants' high confidence ratings reflect a successful understanding of 4-D space and objects. Therefore, the results suggest that these participants were able to form mental representations of 4-D spatial structures through 4-D interactions, and that they applied their representations to the novel 4-D task.

Participants J and K used an alternative solution, although they mentioned an awareness of the strategy of mental 4-D eye-point movement. They focused on one 3-D perspective drawing in the stimulus, and used four lines joining each principal vanishing point and the center of the 3-D screen in 3-D space. They assigned each of the four cells appearing in the 3-D perspective drawing to one end of each axis on the response sheet, while remembering the visual change from a one-point 3-D perspective drawing to a four-point 3-D perspective drawing such that these four lines corresponded to the four principal coordinate axes of the response sheet. Then, they simply allocated the remaining four colored cells to the other ends of the 4-D principal coordinate axes. This strategy is more systematic than the first solution. Participants J and K discovered this solution from a deep understanding of the hypercube's structure and principal vanishing points. They mentioned noticing the approach based on 4-D eye-point control simulation, but chose this solution to reduce the mental workload. In particular, participant K reported that he instantly imagined the entire image of the hypercube while looking at the 3-D perspective drawings and without thinking deeply. Although the exact nature of the structure of their mental representations is not clear, we are probably justified in thinking that participants J and K were able to acquire 4-D spatial representations.

In contrast to these successful participants, participants B, H, and I could not reconstruct the hypercube. We consider some possible scenarios for this failure. First, these participants failed to expand their spatial representations from three to four dimensions. The results of participants B and H fit this scenario, as they reported being able to find the four sets of two cells that faced each other in 4-D space, but could not assemble them correctly because they failed to determine whether the reconstructed coloring pattern was correct or if it was a coloring pattern corresponding to the 4-D mirror image of a stimulus hypercube. It appears that they tried to compare the four 3-D perspective drawings in 3-D space. Second, the participants completely failed to acquire 4-D spatial representations. This scenario is consistent with participant I's result. She reported trying to imagine the 4-D structure of the hypercube, but was unable to. Based on these comments, we conclude that these three participants were unable to apply their experience to the mental simulation and manipulation of 4-D space and objects.

VI. General Discussion

In this paper, we have investigated whether humans are capable of acquiring 4-D spatial representations. Twelve participants interactively observed and learned about 4-D space and a hypercube with the use of an interactive system that utilizes the principal vanishing points as the interface for 4-D eye-point control. Their performance on two different 4-D spatial ability tests was then examined. Seven participants (A, C, D, G, J, K, and L) demonstrated the ability to imagine themselves at a different perspective in 4-D space with respect to the observation object, to update their understanding of an object while moving in 4-D space, and, having observed the object, to understand the relationship between multiple views of the object. Although the exact nature of the 4-D spatial representations is still not known, the present results suggest, at the very least, that mental representations of 4-D space and objects can be formed from experience and interactions in a 4-D environment.

A. 4-D Spatial Representations Versus Reinforcement Learning

The primary concern about our experimental results is whether the participants' performance can be explained by the use of true 4-D spatial representations or, instead, by other low-level solutions. Although the participants' solutions varied, we can rule out some low-level solutions.

First, the tasks in the experiments cannot be solved with mathematical solutions. The participants were theoretically able to estimate their 4-D position from the principal vanishing points, as we showed in our interactive system in Experiment 1. They were also able to estimate spatial relations of multiple 3-D perspective drawings by calculating the 4-D coordinates of the hypercube vertices based on triangulation, as in Experiment 2. None of the participants reported using this strategy in both experiments. Furthermore, none of the participants had prior knowledge of the geometry of 4-D solids. Even if they did, the relevant calculations would be too difficult to perform in real time during the experiment.

Second, the successful participants did not rely on a trial-and-error approach or chance. The evidence is that, in Experiment 1, the participants' operations on the principal vanishing points and their trajectories in 4-D space show no signs of relying on these approaches. Furthermore, the participants' scores in Experiment 2 cannot be explained by chance.

TABLE IV
CORRELATION FACTORS BETWEEN THE EXPERIMENTAL RESULTS AND THE PRELIMINARY TESTS

Test items	Experiment 1		Experiment 2
	First stage	Second stage	
MRT	0.249	−0.017	−0.130
CCT	0.306	0.072	−0.243
SBSOD	0.464	0.290	−0.110
VISQ	0.760	0.520	0.405

Third, the performance of the successful participants cannot be explained by the use of low-level solutions akin to reinforcement learning, which does not cover 4-D spatial representations. There is a possibility that the participants remembered the whole sequence of 3-D perspective drawings of the observation object and traced one perspective to the next, which was remembered as being closer to the target checkpoint. Because the participants interacted with the same hypercube in both learning task and Experiment 1, this explanation could fit the results of Experiment 1 but not those of Experiment 2. As described in Section V, the participants saw only four 3-D perspective drawings of a hypercube, which had a new coloring pattern, in this test. Thus, it would not be possible to rely on the memory of a sequence of 3-D perspective drawings. Rather, the participants were required to interpolate intermediate views of the given 3-D perspective drawings of the hypercube from different perspectives in 4-D space while taking the spatial structure of the hypercube into account. In other words, the experiments did not require the participants to remember all perspectives of the hypercube, but to learn general rules and operations to help them make 4-D judgments on unfamiliar problems. Therefore, combining the results of both experiments, a reasonable interpretation of the present results is that the successful participants developed 4-D spatial representations that help in perspective taking, navigational skills, and mental spatial transformations in 4-D space and which go beyond simple reinforcement learning.

B. Individual Variation Among the Participants

There is some apparent variation among the participants, both in the amount of 4-D experience needed to learn 4-D spatial representations and in their performance level on the 4-D spatial ability tests. We know it is difficult to determine the actual cause for this individual variation because of the small number of participants. However, it is important to study this point for future 4-D spatial cognitive research.

One possible interpretation of the individual variation is that the participants' intrinsic spatial ability, spatial confidence, and focus of attention on the tests varied and affected their 4-D spatial learning. We can verify this interpretation by calculating the correlation factors between the scores on the 4-D spatial ability tests and the scores on the preliminary tests. Table IV presents the correlation factors. Against our expectations, the scores on the 4-D spatial ability tests and the 3-D spatial ability tests are not significantly correlated. In contrast, the hypercube recognition tests and the CCT exhibit a weak negative correlation, even though they differ only in dimension. This suggests that the participants' intrinsic spatial ability has a negligible impact on their development of 4-D spatial representations. Furthermore, this generates the hypothesis that the processing of 4-D visual information and 3-D visual information is related to different mechanisms.

There are positive correlations between the scores on the 4-D spatial ability tests and 3-D spatial confidence. In particular, the correlation between the 4-D spatial ability tests and VISQ is significant. This suggests that self-awareness of good spatial skills encourages the development of 4-D spatial representations. Some recent studies [43], [44] investigating the effect of stereotypical threats on human spatial ability support this suggestion.

There is also a general trend whereby participants from science and engineering courses outperform those from arts courses. Six out of seven participants from science and engineering courses successfully completed the experiments, whereas only one out of five participants from arts courses was able to acquire 4-D spatial representations. Moreover, this arts student had studied science at a high level until fairly recently. As researchers have pointed out, there is some relationship between spatial ability and mathematical aptitude [45], [46]. Thus, it is not unexpected that the participants' intrinsic mathematical knowledge might have positively affected their 4-D learning.

C. Possible Forms of 4-D Spatial Representations

In Section II, we reviewed several previous studies that examine the possibility of acquiring 4-D spatial representations. Because these only demonstrate that humans can acquire 4-D spatial representations, it is important to discuss possible forms of 4-D spatial representations. Certain explanations are consistent with the participants' strategies in the experimental tasks.

One plausible explanation is that the participants formed a cognitive coordinate system consisting of an origin and four directional axes. This type of representation has been demonstrated in many studies on 3-D spatial cognition. Consequently, it is natural that humans might represent 4-D space and objects within a 4-D coordinate system. In particular, in these experiments, because the participants used principal vanishing points, which represent the points at infinity on the 4-D world-coordinate axes, and because they observed a hypercube that was positioned so that each cell was perpendicular to one corresponding 4-D principal coordinate axis and parallel to the other three 4-D principal coordinate axes, there is a high probability that the participants developed and used this type of 4-D spatial representation. In Experiment 1, if the participants formed a 4-D cognitive coordinate system, they would be able to make a rough estimate of the positional relationship between themselves and a target checkpoint in 4-D space, and to guide themselves to the checkpoint position. This type of spatial representation is consistent with both direct flight strategy and relay-points strategy. In Experiment 2, with this coordinate system, they would be able to reconstruct a hypercube by considering its orientation within the 4-D world-coordinate system, which is invariant with respect to the 4-D viewing direction. In particular, the strategy used by participants J and K supports the validity of this explanation.

Although the 4-D cognitive coordinate system seems appropriate as the hypothetical form of the 4-D spatial representations, it remains to be determined whether the four axes are perpendicular to each other in the participants' mental space. As recent spatial cognitive research has shown that the primary form of human spatial knowledge is a graph structure, rather than an absolute map-like coordinate system [8], it is highly likely that the cognitive 4-D coordinate system of 4-D spatial representations does not have strict orthogonal axes. Therefore, the participants are assumed to have performed the experimental tasks based on static mental representations of 4-D space and objects. The question of the orthogonality of the 4-D cognitive coordinate system will be examined in future work.

Another possible explanation is that the participants perceived the hypercube as a 3-D object that changed its appearance according to special transformation rules. It is possible to solve the perspective taking test by memorizing the rules of shape transformations or the sequence of color changes in the 3-D perspective drawings for all patterns of operations on the principal vanishing points. The unsuccessful participants may have relied on this type of representation. For the successful participants, however, this explanation can be ruled out by their actions in Experiment 1. For example, when the participants were at a 4-D position on the x_w-axis, only principal vanishing point vp_x appeared at the center of the 3-D screen. In this case, the participants could not operate principal vanishing point vp_x in the z_s-direction in 3-D space because of the restriction of the interactive system. In other words, in this situation, the possible movement directions were limited in 4-D space. A similar situation occurred when the participants were at a 4-D position on the y_w-axis, but despite this restriction, some successful participants tried to operate the principal vanishing point in the prohibited direction in 3-D space. If the interactive system were improved to allow the participants to perform such an operation, they could have directly approached the target checkpoint in 4-D space. Because they could not experience such 4-D eye-point movement, conceiving this action required the participants to imagine the 4-D spatial relationship between their own position and the target position. Consequently, the fact that the participants attempted this kind of action indicates the successful acquisition of actual 4-D spatial representations.

D. Perspective Taking Versus Object Rotations

Although our interaction technique includes the geometric computation of 4-D eye-point movements, it is difficult to distinguish the 3-D perspective drawings from those obtained by object rotation. This may allow the participants to interpret the 4-D experience in two ways: 1) to imagine eye-point movements in relation to the hypercube or 2) to imagine rotations of the hypercube about a plane or planes passing through the center of the hypercube. Indeed, according to the comments of the seven successful participants, three (C, K, and L) used the former interpretation and four (A, D, G, and J) used the latter. We consider this difference to underlie the embodiment of spatial cognition. Research has shown that perspective taking is related to transformations of the internal representations of the body [47]–[49], whereas object rotations are related to the image of hand motions [50], [51]. In the present experiments, changes in the visual appearance of the hypercube are linked to the participants' hand motion for pick-and-move operations of the principal vanishing points. Although the participants walked around the 3-D virtual space to see the 3-D perspective drawing displayed on the 3-D screen, they stopped at a position from which it was easy to operate the principal vanishing points when moving in 4-D space. Thus, the participants who imagined object rotations may have felt that they were rotating the hypercube by holding and steering a shaft projecting from the side of the hypercube. In contrast, the participants who imagined the 4-D eye-point movements intentionally imagined such actions when they operated the principal vanishing points.

We consider the embodiment to be related to the form of the cognitive coordinate system. Although the present results indicate that the participants acquired 4-D spatial representations in their mental space, the cognitive coordinate system seems to be world-centered or object-centered, as the reference point is located outside the body. A body-centered coordinate system with a reference point inside the body is also necessary to understand 4-D space and objects more intuitively. In particular, when we explore intrinsic 4-D environments, such as a 4-D maze, we need to convert these cognitive coordinate systems into one another. An interaction technique that promotes the acquisition of internal representations of the body extended to 4-D space is an interesting topic for future studies of 4-D spatial cognition.

E. Limitations and Future Work

We offer some notes regarding the limitations of this paper. First, although the results support the possibility of acquiring 4-D spatial representations, our contribution represents one part of the entire scope of 4-D spatial cognition. Using hypercubes and principal vanishing points is too simple to investigate general 4-D spatial representations. The small number of participants made it difficult to interpret the actual cause of the differences among the experimental results. If we were to examine general 4-D spatial representations, we would need more participants to be trained and tested using complex 4-D objects. We consider that the orthogonality and parallelism learned through interaction with a hypercube can form the basis for representations of 4-D space in a cognitive coordinate system. As suggested by the results of the experiments, for some participants, even a hypercube is difficult to understand. General 4-D spatial representations that support 4-D judgments of complex 4-D objects require long-term training after adapting to a simple object.

Second, the scope of this paper was limited to the ability to imagine the visual appearance of an object from different perspectives. Spatial visualization ability in a large-scale 4-D scene was not assessed. In future, a cognitive map of an intricate 4-D environment should be studied, as has been done for 3-D spatial cognition [1], [8], [52]. To acquire knowledge of the environment, people would need to convert their reference spatial representations from the first-person view to the 4-D bird's-eye view. In other words, they would need to convert the cognitive coordinate system from a dynamic one referring to their 4-D position to a static one fixed in 4-D space. To investigate this point, we are currently improving our interactive system to allow 4-D maze exploration in future work.

Third, although the 4-D cognitive coordinate system is one possible 4-D spatial representation, other interaction methods may foster different forms of 4-D spatial representations. For instance, a haptic interface would enhance the recognition of the shape of 4-D objects [53], [54]. The present results suggest that 4-D spatial representations can be acquired through interactions with 4-D objects by principal vanishing point operations, whereas the effects of interface design and interaction style on 4-D spatial representations require future research.

VII. Conclusion

In this paper, we examined the possibility of human 4-D spatial representations through experiments that assess the ability to perform perspective taking, navigation, and mental spatial transformations of objects in 4-D space. The exact nature of the participants' mental representations is not known, but the results suggest that humans are capable of acquiring 4-D spatial representations and using 4-D spatial skills. We therefore succeeded in providing empirical evidence for 4-D spatial representations.

One important aspect of this paper is that it provides a new perspective for research on 4-D spatial representations. Conventional studies on 4-D spatial cognition focus on scenarios in which humans observe a 4-D object in a static condition. In addition, previous research has only considered the ability to comprehend the geometric properties of 4-D objects. In contrast, the present experiments employed free interaction with a 4-D object and examined the ability to manipulate 4-D imagery, which involves cognitive processing, such as prediction, creation, and decision making for spatial visual information. The results suggest that humans are capable of acquiring 4-D spatial representations and that this cognitive processing works properly in 4-D space. This means that human spatial cognition does not have intrinsic dimensionality constraints, even though humans evolved in a 3-D world. Instead, it is thought that human cognitive processing is flexible and can adapt to higher-dimensional space with practice and experience of 4-D space and objects.

Appendix A
Algorithms for 4-D Eye-Point Control via Principal Vanishing Point Operations

As described in Section III, a 3-D perspective drawing of a 4-D object was obtained by converting data defined in the 4-D world-coordinate system $x_w y_w z_w w_w$ to data in the 3-D screen-coordinate system $x_s y_s z_s w_s$ [30], [31]. The homogeneous coordinates V_w including the points at infinity in 4-D space are transformed to the homogeneous coordinates V_s as follows:

$$\begin{aligned} \boldsymbol{V}_s &= [X_s \quad Y_s \quad Z_s \quad W_s \quad v_s] \\ &= [X_w \quad Y_w \quad Z_w \quad W_w \quad v_w]\boldsymbol{T}_v\left(p_f, p_a\right)\boldsymbol{T}_p(k, h, f) \end{aligned} \tag{1}$$

where the transformation matrices $\boldsymbol{T}_v$ and $\boldsymbol{T}_p$ are the 4-D view field transformation matrix and the 4-D perspective transformation matrix, respectively. The 4-D view field transformation matrix $\boldsymbol{T}_v$ is derived from the 4-D eye-point $p_f(x_{p_f}, y_{p_f}, z_{p_f}, w_{p_f})$ and the 4-D observed point $p_a(x_{p_a}, y_{p_a}, z_{p_a}, w_{p_a})$ as follows:

$$\begin{aligned} \boldsymbol{T}_v\left(p_f, p_a\right) = \boldsymbol{T}_t\left(-x_{p_f}, -y_{p_f}, -z_{p_f}, -w_{p_f}\right)\boldsymbol{T}_{xy}(\sin\alpha, \cos\alpha) \\ \times\ \boldsymbol{T}_{yz}(\sin\beta, \cos\beta)\boldsymbol{T}_{xz}(\sin\gamma, \cos\gamma) \end{aligned} \tag{2}$$

where

$$\left.\begin{aligned}
\cos\alpha &= \frac{w_{p_f} - w_{p_a}}{\sqrt{\left(z_{p_f}-z_{p_a}\right)^2+\left(w_{p_f}-w_{p_a}\right)^2}} \\
\sin\alpha &= \frac{z_{p_f} - z_{p_a}}{\sqrt{\left(z_{p_f}-z_{p_a}\right)^2+\left(w_{p_f}-w_{p_a}\right)^2}} \\
\cos\beta &= \frac{\sqrt{\left(z_{p_f}-z_{p_a}\right)^2+\left(w_{p_f}-w_{p_a}\right)^2}}{\sqrt{\left(x_{p_f}-x_{p_a}\right)^2+\left(z_{p_f}-z_{p_a}\right)^2+\left(w_{p_f}-w_{p_a}\right)^2}} \\
\sin\beta &= \frac{x_{p_a} - x_{p_f}}{\sqrt{\left(x_{p_f}-x_{p_a}\right)^2+\left(z_{p_f}-z_{p_a}\right)^2+\left(w_{p_f}-w_{p_a}\right)^2}} \\
\cos\gamma &= \frac{\sqrt{\left(x_{p_f}-x_{p_a}\right)^2+\left(z_{p_f}-z_{p_a}\right)^2+\left(w_{p_f}-w_{p_a}\right)^2}}{\sqrt{\left(x_{p_f}-x_{p_a}\right)^2+\left(y_{p_f}-y_{p_a}\right)^2+\left(z_{p_f}-z_{p_a}\right)^2+\left(w_{p_f}-w_{p_a}\right)^2}} \\
\sin\gamma &= \frac{y_{p_a} - y_{p_f}}{\sqrt{\left(x_{p_f}-x_{p_a}\right)^2+\left(y_{p_f}-y_{p_a}\right)^2+\left(z_{p_f}-z_{p_a}\right)^2+\left(w_{p_f}-w_{p_a}\right)^2}}
\end{aligned}\right\}.$$

The transformation matrix $\boldsymbol{T}_t$ is the 4-D translation matrix, and the transformation matrices $\boldsymbol{T}_{xy}$, $\boldsymbol{T}_{yz}$, and $\boldsymbol{T}_{xz}$ are the 4-D rotation matrices around the xy-, yz-, and xz-planes, respectively. Therefore, the 3-D screen coordinates v_s are obtained from (1) as follows:

$$\begin{aligned} v_s &= (x_s, y_s, z_s, w_s) \\ &= \left(\frac{X_s}{v_s}, \frac{Y_s}{v_s}, \frac{Z_s}{v_s}, \frac{W_s}{v_s}\right). \end{aligned} \tag{3}$$

This algorithm enables various types of 4-D data to be observed from an arbitrary 4-D viewing direction for an arbitrary 4-D eye-point. Moreover, by changing the parameters k, h, and f of the 4-D perspective transformation matrix $\boldsymbol{T}_p$, we can visualize 4-D data not only with various 4-D viewing fields, but also with projection methods, such as a perspective projection, parallel projection, or slice operation.

Now, we consider the principal vanishing points in 3-D space. The points at infinity in the x_w-, y_w-, z_w-, and w_w-directions are represented as $V_{x_w}\overline{(1, 0, 0, 0, 0)}$, $V_{y_w}\overline{(0, 1, 0, 0, 0)}$, $V_{z_w}\overline{(0, 0, 1, 0, 0)}$, and $V_{w_w}\overline{(0, 0, 0, 1, 0)}$, respectively. Substituting these points into (1), we can obtain

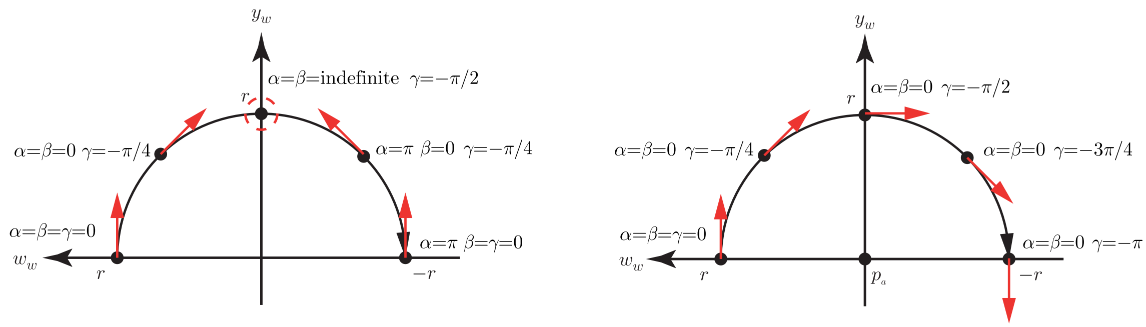


Fig. 10. Difference in attitude of the 4-D eye-point caused by modification of the implementation. (a) Original visualization algorithm. (b) Our implementation.

the principal vanishing points vp_x, vp_y, vp_z, and vp_w on the 3-D screen from (3) as follows:

$$\begin{aligned}
vp_x &= \left(x_{vp_x}, y_{vp_x}, z_{vp_x}\right)\\
&= \left(\frac{h}{k}\frac{1}{\tan\beta\cos\gamma}, -\frac{h}{k}\tan\gamma, 0\right)\\
vp_y &= \left(x_{vp_y}, y_{vp_y}, z_{vp_y}\right)\\
&= \left(0, \frac{h}{k}\frac{1}{\tan\gamma}, 0\right)\\
vp_z &= \left(x_{vp_z}, y_{vp_z}, z_{vp_z}\right)\\
&= \left(-\frac{h}{k}\frac{\tan\beta}{\cos\gamma}, -\frac{h}{k}\tan\gamma, -\frac{h}{k}\frac{1}{\tan\alpha\cos\beta\cos\gamma}\right)\\
vp_w &= \left(x_{vp_w}, y_{vp_w}, z_{vp_w}\right)\\
&= \left(-\frac{h}{k}\frac{\tan\beta}{\cos\gamma}, -\frac{h}{k}\tan\gamma, \frac{h}{k}\frac{\tan\alpha}{\cos\beta\cos\gamma}\right)
\end{aligned}\tag{4}$$

where α, β, and γ are the parameters of the 4-D viewing direction for the 4-D view field transformation.

From the converse relation of (4), it is assumed that we can derive the parameters α, β, and γ if some principal vanishing points are given in 3-D space. Based on this assumption, we constructed an interaction algorithm that made position changes of the principal vanishing points in 3-D space correspond to movement of the 4-D eye-point in 4-D space [16]. This algorithm is composed of two processing steps. The first estimates the positions of the principal vanishing points in 3-D space in accordance with user operation. The second estimates the parameters of the 4-D viewing direction in 4-D space using the principal vanishing points in 3-D space to determine the 4-D eye-point.

We now discuss the first processing step. When one principal vanishing point is picked and moved by a user in 3-D space, the other three principal vanishing points should be automatically allocated to the corresponding positions that satisfy their geometric positional relationship in 3-D space. For instance, suppose the principal vanishing points vp_{x_b}, vp_{y_b}, vp_{z_b}, and vp_{w_b} are displayed in 3-D space at a 4-D eye-point p_{f_b}. When the principal vanishing point vp_{w_b} moves to vp_w through user operations, (4) allows us to estimate the other three vanishing points vp_x, vp_y, and vp_z using the position of the operated principal vanishing point vp_w as follows:

$$\begin{aligned}
vp_x &= \left(x_{vp_x}, y_{vp_x}, z_{vp_x}\right)\\
&= \left(-\frac{1}{x_{vp_w}}\left\{\left(\frac{h}{k}\right)^2 + {y_{vp_w}}^2\right\}, y_{vp_w}, 0\right)\\
vp_y &= \left(x_{vp_y}, y_{vp_y}, z_{vp_y}\right)\\
&= \left(0, -\frac{1}{y_{vp_w}}\left(\frac{h}{k}\right)^2, 0\right)\\
vp_z &= \left(x_{vp_z}, y_{vp_z}, z_{vp_z}\right)\\
&= \left(x_{vp_w}, y_{vp_w}, -\frac{1}{z_{vp_w}}\left\{\left(\frac{h}{k}\right)^2 + {x_{vp_w}}^2 + {y_{vp_w}}^2\right\}\right).
\end{aligned}\tag{5}$$

If another principal vanishing point is operated on, or if any principal vanishing points are not displayed in 3-D space, the estimation of the principal vanishing points can be achieved in the same manner as in (5).

Next, we discuss the second processing step. For 4-D eye-point control, we consider the movement of 4-D eye-point p_f along a 4-D spherical surface with radius r. The 4-D spherical surface is centered at the origin of the 4-D world-coordinate system. The 4-D observed point p_a is fixed at the origin of the 4-D world-coordinate system. Parameters α, β, and γ of the 4-D viewing direction are derived from (4) as follows:

$$\begin{aligned}
\alpha &= \tan^{-1}\frac{z_{vp_w}}{\sqrt{-z_{vp_z}z_{vp_w}}}\\
\beta &= \tan^{-1}\frac{-x_{vp_w}}{\sqrt{-x_{vp_x}x_{vp_w}}}\\
\gamma &= \tan^{-1}\frac{-y_{vp_w}}{\sqrt{-y_{vp_y}y_{vp_w}}}.
\end{aligned}\tag{6}$$

Substituting the coordinate values of the principal vanishing points, for example, the coordinate values of (5), into (6), we can determine the corresponding parameters α, β, and γ. Finally, the 4-D eye-point p_f is computed with the following equation:

$$\begin{aligned}
p_f &= \begin{bmatrix} x_{p_f} & y_{p_f} & z_{p_f} & w_{p_f} & 1\end{bmatrix}\\
&= [0\ \ 0\ \ 0\ \ r\ \ 1]\boldsymbol{T}_{xz}^{-1}(\gamma)\boldsymbol{T}_{yz}^{-1}(\beta)\boldsymbol{T}_{xy}^{-1}(\alpha)
\end{aligned}\tag{7}$$

where the transformation matrices $\boldsymbol{T}_{xz}$, $\boldsymbol{T}_{yz}$, and $\boldsymbol{T}_{xy}$ represent the 4-D rotation matrices around the xz-, yz-, and xy-planes, respectively.

Appendix B

Implementation of the Visualization Algorithm and the Interaction Algorithm

As mentioned in Section III, we apply the parameters α, β, and γ in (6) to the 4-D view field transformation matrix $\boldsymbol{T}_v$ in (1), instead of introducing the 4-D eye-point p_f in (7) and the 4-D observed point p_a into (1). That is, the 4-D view field transformation matrix $\boldsymbol{T}_v$ of (2) is arranged and represented as follows:

$$\begin{aligned}&\boldsymbol{T}_v\left(p_f, \alpha, \beta, \gamma\right)\\&\quad= \boldsymbol{T}_t\left(-x_{p_f}, -y_{p_f}, -z_{p_f}, -w_{p_f}\right)\boldsymbol{T}_{xy}(\alpha)\boldsymbol{T}_{yz}(\beta)\boldsymbol{T}_{xz}(\gamma). \qquad (8)\end{aligned}$$

In this appendix, we explain the effect of this modification on the implementation with a concrete example to better clarify the 4-D interaction.

Suppose that the 4-D eye-point starts at the 4-D position $(0, 0, 0, r)$ on the w_w-axis and moves to $(0, 0, 0, -r)$ via a zenith of the 4-D spherical surface $(0, r, 0, 0)$ on the y_w-axis, where $r(> 0)$ represents the radius of the 4-D spherical surface. When the 4-D eye-point is in the 4-D half-space, which is defined as the 4-D region satisfying $0 \leq w_w$, the modification of the implementation does not affect the visualization and the interaction because the 4-D view field transformation matrix $\boldsymbol{T}_v$ of (2) coincides with that of (8). The effect of the modification becomes apparent once the 4-D eye-point arrives at the zenith and enters the other 4-D half-space, which is defined as the 4-D region satisfying $w_w < 0$. This is the attitude of the moving 4-D eye-point.

Fig. 10 shows the difference in the 4-D attitude brought about by the modification of the implementation. For simplification, we only describe the $y_w w_w$-plane, which includes the movement path of the 4-D eye-point, and focus on five positions along the movement path. The attitude of the 4-D eye-point at each position is expressed by the 4-D upper direction of the 4-D eye-point, represented by a red arrow. In addition, in the figure, we include the values α, β, and γ of the 4-D viewing direction for the corresponding method.

When we visualize 4-D space and objects according to the original method using the 4-D view field transformation matrix $\boldsymbol{T}_v$ of (2), which is derived from the 4-D eye-point p_f of (7) and the 4-D observed point p_a, the values α, β, and γ of the 4-D viewing direction are in the range $-\pi \leq \alpha \leq \pi$ and $-\pi/2 \leq \beta, \gamma \leq \pi/2$. In this case, the 4-D upper direction is indefinite at the zenith of the 4-D spherical surface and is reversed by 180° before and after the 4-D eye-point passes through the zenith, as shown in Fig. 10(a). The continuity of 4-D observation is therefore lost, and the interaction becomes unnatural at the zenith.

To remove this irregularity, we determine the parameters α, β, and γ of the 4-D viewing direction of (6) in the range $-\pi \leq \alpha, \beta, \gamma \leq \pi$, depending on the movement history of the 4-D eye-point, so that the attitude of the 4-D eye-point is maintained before and after it passes through the zenith, as shown in Fig. 10(b). Then, we do not generate the 4-D view field transformation matrix $\boldsymbol{T}_v$ of (2), but that of (8). As a result, we avoid the discontinuity in 4-D observation and provide 4-D interaction in a natural style.

References

[1] E. C. Tolman, "Cognitive maps in rats and men," *Psychol. Rev.*, vol. 55, no. 4, pp. 189–208, Jul. 1948.

[2] J. Piaget and B. Inhelder, *The Child's Conception of Space*. New York, NY, USA: Norton, May 1967.

[3] R. N. Shepard and J. Metzler, "Mental rotation of three-dimensional objects," *Science*, vol. 171, no. 3972, pp. 701–703, Feb. 1971.

[4] M. A. Just and P. A. Carpenter, "Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability," *Psychol. Rev.*, vol. 92, no. 2, pp. 137–172, Apr. 1985.

[5] T. J. Branoff, "Spatial visualization measurement: A modification of the purdue spatial visualization test—Visualization of rotations," *Eng. Design Graph. J.*, vol. 64, no. 2, pp. 14–22, 2000.

[6] M. Kozhevnikov and M. Hegarty, "A dissociation between object manipulation spatial ability and spatial orientation ability," *Memory Cogn.*, vol. 29, no. 5, pp. 745–756, Jul. 2001.

[7] M. Hegarty and D. Waller, "A dissociation between mental rotation and perspective-taking spatial abilities," *Intelligence*, vol. 32, no. 2, pp. 175–191, Mar./Apr. 2004.

[8] E. R. Charstil and W. H. Warren, "From cognitive maps to cognitive graphs," *PLoS ONE*, vol. 9, no. 11, Nov. 2014, Art. no. e112544.

[9] P. J. Davis, R. Hersh, and E. A. Marchisotto, "Four dimensional intuition," in *The Mathematical Experience, Study Edition*. Cambridge, MA, USA: Birkhäuser, 1995, pp. 442–447.

[10] T. N. Aflalo and M. S. Graziano, "Four-dimensional spatial reasoning in humans," *J. Exp. Psychol. Human Perception Perform.*, vol. 34, no. 5, pp. 1066–1077, Oct. 2008.

[11] M. S. Ambinder, R. F. Wang, J. A. Crowell, G. K. Francis, and P. Brinkmann, "Human four-dimensional spatial intuition in virtual reality," *Psychonomic Bull. Rev.*, vol. 16, no. 5, pp. 818–823, Oct. 2009.

[12] R. F. Wang, "A case study on human learning of four-dimensional objects in virtual reality: Passive exploration and display techniques," in *Proc. 4th Int. Conf. Front. Comput. Sci. Technol.*, Dec. 2009, pp. 519–523.

[13] R. F. Wang, "Can humans form four-dimensional spatial representations?" in *Space in Mind: Concepts for Spatial Learning and Education*. Cambridge, MA, USA: MIT Press, Nov. 2014, pp. 119–136.

[14] R. F. Wang, "Human four-dimensional spatial judgments of hyper-volume," *Spatial Cogn. Comput. Interdiscipl. J.*, vol. 14, no. 2, pp. 91–113, Jan. 2014.

[15] T. Miwa, Y. Sakai, and S. Hashimoto, "4-D spatial perception established through hypercube recognition tasks using interactive visualization system with 3-D screen," in *Proc. ACM SIGGRAPH Symp. Appl. Perception*, Tübingen, Germany, Sep. 2015, pp. 75–82.

[16] T. Miwa, Y. Sakai, and S. Hashimoto, "Four-dimensional eye-point control by principal vanishing points operation and its evaluation," *J. Soc. Art Sci.*, vol. 12, no. 4, pp. 162–174, Dec. 2013.

[17] T. Miwa, Y. Sakai, and S. Hashimoto, "Four-dimensional viewing direction control by principal vanishing points operation and its application to four-dimensional fly-through experience," in *Proc. 25th Aust. Comput. Human Interaction Conf. Augmentation Appl. Innov. Collaboration*, Adelaide, SA, Australia, Nov. 2013, pp. 95–104.

[18] T. Miwa, Y. Sakai, and S. Hashimoto, "Interactive four-dimensional space exploration using viewing direction control based on principal vanishing points operation," in *Computer-Human Interaction. Cognitive Effects of Spatial Interaction, Learning, and Ability* (LNCS 8433). Cham, Switzerland: Springer, Apr. 2015, pp. 21–46.

[19] A. K. Dewdney, "COMPUTER RECREATIONS: A program for rotating hypercubes induces four-dimensional dementia," *Sci. Amer.*, vol. 254, pp. 8–13, Apr. 1986.

[20] A. J. Hanson, "Computer graphics beyond the third dimension," in *Course Notes for SIGGRAPH '98*. Orlando, FL, USA, Jul. 1998.

[21] B. Hausmann and H.-P. Seidel, "Visualization of regular polytopes in three and four dimensions," *Comput. Graph. Forum*, vol. 13, no. 3, pp. 305–316, Aug. 1994.

[22] S. R. Hollasch, "Four-space visualization of 4D objects," M.S. thesis, Arizona State Univ., Tempe, AZ, USA, Aug. 1991. [Online]. Available: http://steve.hollasch.net/thesis/

[23] K. Miyazaki and K. Ishihara, *Four-Dimensional Graphics*. Tokyo, Japan: Asakura-Shoten, Sep. 1989.

[24] K. Miyazaki, M. Ishii, and S. Yamaguchi, *Science of Higher-Dimensional Shapes and Symmetry*. Kyoto, Japan: Kyoto Univ. Press, Feb. 2005.

[25] N. Neophytou and K. Mueller, "Space-time points: 4D splatting on efficient grids," in *Proc. IEEE Symp. Vol. Vis. Graph.*, Boston, MA, USA, Oct. 2002, pp. 97–106.

[26] J. Woodring, C. Wang, and H.-W. Shen, “High dimensional direct rendering of time-varying volumetric data,” in *Proc. 14th IEEE Vis.*, Oct. 2003, pp. 417–424.

[27] M. D’Zmura, P. Colantoni, and G. D. Seyranian, “Virtual environments with four or more spatial dimensions,” *Presence*, vol. 9, no. 6, pp. 616–631, Dec. 2000.

[28] X. Yan, C.-W. Fu, and A. J. Hanson, “Multitouching the fourth dimension,” *IEEE Comput.*, vol. 45, no. 9, pp. 80–88, Sep. 2012.

[29] J. C. Aguilera, “Virtual reality and the unfolding of higher dimensions,” in *Stereoscopic Displays and Virtual Reality Systems XIII* (Proceedings of SPIE), vol. 6055. Bellingham, WA, USA: SPIE, Jan. 2006, pp. 605–612.

[30] Y. Sakai and S. Hashimoto, “Interactive four-dimensional space visualization using five-dimensional homogeneous processing for intuitive understanding,” *J. Inst. Image Inf. Television Eng.*, vol. 60, no. 10, pp. 1630–1647, Oct. 2006.

[31] Y. Sakai and S. Hashimoto, “Four-dimensional space visualization with four-dimensional viewing field control,” *Bull. Soc. Sci. Form*, vol. 21, no. 3, pp. 274–284, Mar. 2007.

[32] Y. Sakai and S. Hashimoto, “Four-dimensional space-time visualization for understanding three-dimensional motion,” *J. Inst. Image Electron. Eng. Japan*, vol. 36, no. 4, pp. 371–381, Jul. 2007.

[33] Y. Sakai and S. Hashimoto, “Four-dimensional mathematical data visualization via ‘embodied four-dimensional space display system,”’ *Forma*, vol. 26, no. 1, pp. 11–18, May 2011.

[34] Y. Sakai and S. Hashimoto, “Four-dimensional geometric element definitions and interferences via five-dimensional homogeneous processing,” *J. Vis.*, vol. 14, no. 2, pp. 129–139, Jun. 2011.

[35] M. Hegarty, “Mechanical reasoning by mental simulation,” *Trends Cogn. Sci.*, vol. 8, no. 6, pp. 280–285, Jun. 2004.

[36] D. L. Schwartz and J. B. Black, “Shuttling between depictive models and abstract rules: Induction and fallback,” *Cogn. Sci.*, vol. 20, no. 4, pp. 457–497, Oct. 1996.

[37] M. Stieff, “Mental rotation and diagrammatic reasoning in science,” *Learn. Instruct.*, vol. 17, no. 2, pp. 219–234, Apr. 2007.

[38] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics: Principles and Practice*, 2nd ed. Reading, MA, USA: Addison-Wesley, Aug. 1995.

[39] K. Kondo, F. Kimura, and T. Tajima, “Estimation of a point of view with perspective drawing and the application,” *Trans. Inf. Process. Soc. Japan*, vol. 29, no. 7, pp. 686–693, 1988.

[40] M. Peters *et al.*, “A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance,” *Brain Cogn.*, vol. 28, no. 1, pp. 39–58, Jun. 1995.

[41] M. Hegarty, A. E. Richardson, D. R. Montello, K. Lovelace, and I. Subbiah, “Development of a self-report measure of environmental spatial ability,” *Intelligence*, vol. 30, no. 5, pp. 425–447, Sep./Oct. 2002.

[42] M. Kawahara and K. Matsuoka, “Development of the visual imagery style questionnaire,” *Jpn. J. Mental Imagery*, vol. 7, no. 1, pp. 19–31, 2009.

[43] M. S. McGlone and J. Aronson, “Stereotype threat, identity salience, and spatial reasoning,” *J. Appl. Develop. Psychol.*, vol. 27, no. 5, pp. 486–493, Sep./Oct. 2006.

[44] Z. Estes and S. Felker, “Confidence mediates the sex difference in mental rotation performance,” *Archives Sexual Behav.*, vol. 41, no. 3, pp. 557–570, Jun. 2012.

[45] J. Wai, D. Lubinski, and C. P. Benbow, “Spatial ability for stem domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance,” *J. Edu. Psychol.*, vol. 101, no. 4, pp. 817–835, Nov. 2009.

[46] Y.-L. Chenga and K. S. Mixa, “Spatial training improves children’s mathematics ability,” *J. Cogn. Develop.*, vol. 15, no. 1, pp. 2–11, Jan. 2014.

[47] C. C. Presson and D. R. Montello, “Updating after rotational and translational body movements: Coordinate structure of perspective space,” *Perception*, vol. 23, no. 12, pp. 1447–1455, Feb. 1994.

[48] O. Blanke *et al.*, “Linking out-of-body experience and self processing to mental own-body imagery at the temporoparietal junction,” *J. Neurosci.*, vol. 25, no. 3, pp. 550–557, Jan. 2005.

[49] M. Kozhevnikov, M. A. Motes, B. Rasch, and O. Blajenkova, “Perspective-taking vs. mental rotation transformations and how they predict spatial navigation performance,” *Appl. Cogn. Psychol.*, vol. 20, no. 3, pp. 397–417, Apr. 2006.

[50] A. Wohlschläger, M. Gattis, and H. Bekkering, “Action generation and action perception in imitation: An instance of the ideomotor principle,” *Philosoph. Trans. Roy. Soc. London B Biol. Sci.*, vol. 358, no. 1431, pp. 501–515, Feb. 2003.

[51] M. Wraga, W. L. Thompson, N. M. Alpert, and S. M. Kosslyn, “Implicit transfer of motor strategies in mental rotation,” *Brain Cogn.*, vol. 52, no. 2, pp. 135–143, Jul. 2003.

[52] K. Lynch, *The Image of the City*. Cambridge, MA, USA: MIT Press, Jun. 1960.

[53] A. J. Hanson and H. Zhang, “Multimodal exploration of the fourth dimension,” in *Proc. IEEE Vis.*, Minneapolis, MN, USA, Oct. 2005, pp. 263–270.

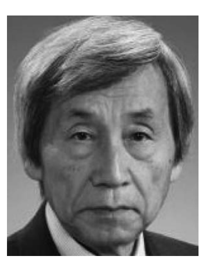
[54] H. Zhang and A. J. Hanson, “Shadow-driven 4D haptic visualization,” *IEEE Trans. Vis. Comput. Graph.*, vol. 13, no. 6, pp. 1688–1695, Nov./Dec. 2007.

Takanobu Miwa received the B.Eng. and M.Eng. degrees in pure and applied physics from Waseda University, Tokyo, Japan, in 2010 and 2012, respectively. From 2012 to 2015, he was a Ph.D. student in the Graduate School of Science and Engineering, Waseda University.

From 2010 to 2015, he was a Part-Time Lecturer with Waseda Jitsugyo Gakko, Tokyo. Since 2015, he has been a Research Associate with the Faculty of Science and Engineering, Waseda University. His current research interests include scientific visualization, cognitive science, human–computer interaction, and image processing.

Yukihito Sakai received the B.E. and M.E. degrees in mechanical engineering, and the Ph.D. degree in pure and applied physics from Waseda University, Tokyo, Japan, in 1999, 2001, and 2008, respectively.

From 2001 to 2004, he was a Systems Engineer with Toshiba Corporation, Tokyo. From 2004 to 2007, he was a Visiting Research Associate with the Faculty of Science and Engineering, Waseda University, where he was a Research Associate, from 2007 to 2009. From 2009 to 2012, he was an Assistant Professor with the Faculty of information Sciences and Arts, Toyo University, Tokyo. Since 2012, he has been an Associate Professor with the Faculty of Engineering, Fukuoka University, Fukuoka, Japan. His current research interests include graphic science, computer graphics, and interactive systems, in particular, 4-D graphics and 4-D interactions.



Shuji Hashimoto (M’03) received the B.S., M.S., and Dr.Eng. degrees in pure and applied physics from Waseda University, Tokyo, Japan, in 1970, 1973, and 1977, respectively.

Since 2000, he was the Director of the Humanoid Robotics Institute, Waseda University, where he is a Professor with the Department of Applied Physics. His current research interests include “KANSEI” information processing, image processing, meta-algorithm, artificial intelligence, and robotics.

Dr. Hashimoto is a member of major academic societies, including the Institute of Electronics, Information and Communication Engineers, Information Processing Society of Japan, and the Robotics Society of Japan.