Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Optimal camera angles in mental imaging of three-dimensional structures from plane figures and the effects of depth cues

Koji Kashihara
Nagoya Institute of Technology (Nagoya Univ.),
Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan
kashihara.koji@nitech.ac.jp (kashihara@cog.human.nagoya-u.ac.jp)

***Abstract*—In order to evaluate the specific brain activities related to the recognition of a three-dimensional (3D) structure, the construction of 3D images in the brain from 2D figures and the ability to match the mental image to an actual 3D object were investigated. The results demonstrated that the effects of depth cues (shade and colors) in the 3D images facilitated the cognitive process in human-computer interfaces and that there existed optimal presentation angles for the matching phase; these findings may indicate effective methods in a field requiring quick judgment based on high cognitive functions.**

Keywords—3D image, human-computer interface, reaction time, depth cues, mental rotation

I. INTRODUCTION

There are some cases when one must create a mental image of a three-dimensional (3D) structure from 2D figures [1]. For example, computer aided design (CAD) or clinical assessment by MRI or CT data require the correct construction of a 3D image in the brain from 2D figures. In these cases, the presentation method of a 3D object may affect the efficiency of the cognitive process. For example, one cannot clearly grasp the depth of a 3D object only from the front view [2]; however, when the object is shown at various angles, one receives information about depth and it becomes easier to image the 3D structure. Even in the 3D image in the brain, depth cues might be used to understand the 3D structure, when mentally rotating the object.

The mental rotation task is well known in the investigation of the cognitive process in the operation of mental images [3, 4]; the 3D figures compared, including mirror images, have the same angles in the face and depth axes at the angular difference of zero degrees, regardless of various camera angles. There is also some evidence that optimal angles for the 3D object exist [5, 6]. However, the optimal angles in the 3D object, compared with the 3D image constructed from 2D figures in the brain, have not been demonstrated in detail.

The purpose of this study, therefore, is to investigate effective methods for 3D presentation, using mental images constructed from 2D figures. First, it was tested whether depth cues can increase the speed of the mental construction of 3D images from 2D figures. In particular, the effects of shadow and figure-ground segregation [1] using color difference on cognitive speed in constructing 3D images were evaluated. Next, it was evaluated how view angles affected the construction of 3D images in the brain. Furthermore, the optimal 3D viewpoints and the actual 3D angles in the brain were explored, using subjective tests, in order to compare with the actual cognitive speed.

II. METHODS

A. Subjects

The subjects were ten healthy male volunteers (mean age, 23.9 ± 4.0 years). Informed consent was obtained from all subjects after a complete description of the experiment. All subjects had normal or corrected-to-normal vision. The experimenter verified the condition of the subjects before the experiment and confirmed that they were healthy.

B. Stimuli

Figure 1 shows the four kinds of 3D figures used in the present study; the number of cubes was five. The four mirror images were also used. There are two types of rotation of a 3D object; the elevation and azimuth of a given viewpoint correspond to the x and y axes (Fig. 2A). In this experiment, the 3D figure was shown at various angles from 15 to 75 degrees in increments of 15 degrees: (Elevation, Azimuth) = [(15, 30, 45, 60, or 75), (15, 30, 45, 60, or 75)]. In addition, the front views were set to 16 variants by the rotation of the axes in each 3D figure (total = 4*2*16 variants).

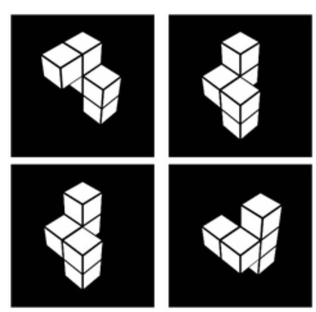

Figure 1. Four examples of 3D figures. The view angles were set to 45 and 45 degrees in azimuth and elevation.

Two-dimensional figures consisted of parts of the 3D figures used (Fig. 2B). This method of presentation is generally used for CAD design or analysis systems for MRI (or CT) scans.

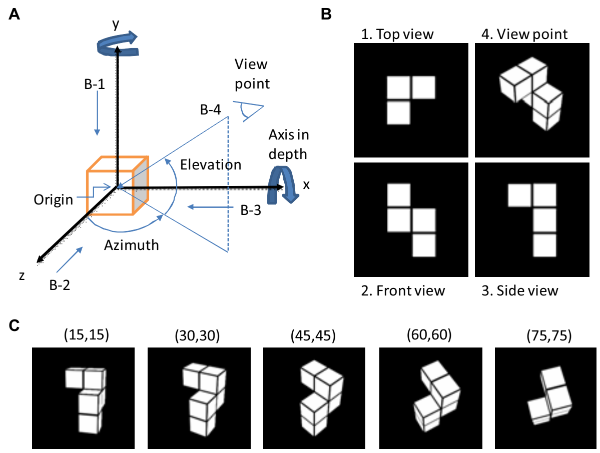

Figure 2. A: the definition of viewpoint and angles of a 3D object. B: the place used for presenting 2D and 3D pictures on a screen. C: Five typical examples of all 25 types of angles. ( ) shows the combination of the view angles in elevation and azimuth.

The visual stimuli were presented on a screen placed in front of the subjects at the same height as their eyes. The distance from the stimuli to the center of the eyes was set at 2 m. The size of the displayed stimuli was no more than 15 cm (per picture) as vertical and horizontal lines. The squares for the simple reaction task were 5 cm per line.

C. *Tasks*

Five types of tasks were performed: 1) optimal 3D angle, 2) angle of 3D image, 3) simple reaction to colored squares, 4) 3D image from colored 2D figures (two of the subjects did not perform this task), and 5) 3D image from non-colored 2D figures.

1) Optimal 3D angle

Subjects were to seek the optimal angle from which one can easily view the entire structure of a 3D object. The viewpoint angles ranged from 15 to 75 degrees in both azimuth and elevation, in increments of 15 degrees (Fig. 2C). The subjects set the optimal angles at which they felt it was easiest to recognize the overall structure of the 3D object presented. It was evaluated whether the optimal angles in this task were related to faster response of in task 5).

2) Angle of 3D image

Subjects mentally imaged a 3D structure from three 2D figures showing the top, front, and side views of 3D. The subjects then set the angles of a newly-presented 3D figure to correspond to the mental image they had produced. The view angles were 15, 30, 45, 60, and 75 degrees in both azimuth and elevation (Fig. 2C). The relationship between the selected angles of 3D image in this task and the results of task 5) were investigated.

3) Simple reaction to colored squares

To evaluate response time to colored squares, a simple reaction task was performed in which subjects pressed a mouse button. The used color squares were 6 types: white, light gray, deep gray, red, green, and blue.

4) 3D image from colored 2D figures

The effect of depth cues was evaluated using the reaction time required to mentally image a 3D object and compare the image with an actual object. Subjects attempted to image the 3D structure from 2D figures as accurately and quickly as possible. The depth cues in the 2D figures were supplied by shadow (white, light gray and dark gray; the shadow condition) and color (red, green, and blue; the color condition). The nearest surface from the viewpoint was white (or red), the surface one block deeper was light gray (or green), and the surface two blocks deeper was dark gray (or blue). Two-dimensional figures in white only were used as a control.

Next, subjects compared the imaged 3D figure and the newly presented one and indicated “correct” or “incorrect”. The 3D view angles were the combinations of (30, 30), (45, 45), and (60, 60) degrees in both azimuth and elevation. Subjects also filled out a questionnaire to evaluate the subjective effect of shadowed and colored figures on reaction time. The scale had 7 levels in increments of 1: “bad” = 1, “not significant” = 4, and “good” = 7.

5) 3D image from non-colored 2D figures

This task is the almost same as the task 4) above. The different points were as follows. The color of all 2D figures was white, i.e., control condition in the task 4). The view angles were set to 15, 30, 45, 60, and 75 degrees in both azimuth and elevation (i.e., 25 combinations).

D. *Procedures*

The five tasks performed were ordered as follows for all subjects: tasks 1) through 5).

1) Optimal 3D angle

A 3D picture was initially presented on a screen. Subjects selected the optimal viewpoint for understanding the overall structure of the 3D figures. They pressed a button after selecting the optimal angles of each 3D figure, and next trial was immediately started. Thirty trials (randomly selected from 3D dataset) were performed for each subject. The time for this task was not limited.

2) Angle of 3D image

Two-dimensional pictures showing the top, front, and side views of a 3D figure were initially presented on a screen. Subjects pressed a button after constructing an accurate 3D image from the 2D figures without an actual 3D figure; they then selected the optimal angles for a newly presented 3D figure. Thirty trials (randomly selected from the dataset) were performed consecutively for each subject. The time for this task was not limited. Subjects were instructed to familiarize themselves with the figures and the set angles, because they were to be used in the following tasks.

3) Simple reaction to colored squares

A "Start" stimulus was shown at the center of a screen for 500 ms, and the interval was set at 600-1200 ms. Subjects clicked the button in reaction to a square stimulus (300 ms) as quickly as possible. The squares shown were the six colors that were used for the following task 4): white, light gray, and dark gray in the shadow condition, and red, green, and blue in the color condition, with black as a background color.

The subjects performed a simple reaction task to familiarize themselves with pressing the button while looking at the screen (10 trials); they then performed the actual task (60 trials) in addition to the initial 5 trials. The colored squares were randomly ordered, and the total number of appearances for each colored square was 10.

4) *3D image from colored 2D figures*

In each trial, a "Start" stimulus was shown on a screen for 500 ms. After a jittering period of 600-1200 ms, three 2D figures showing the top (0 and 90 degrees in azimuth and elevation), front (0 and 0 degrees), and side (90 and 0 degrees) views of a 3D figure (Fig. 2A) were pseudo-randomly presented from the stimulus sets. Either the control, shadow, or color condition in the 2D figures was randomly selected. The subjects were required to mentally image the 3D structure from the 2D figures presented on a screen (Fig. 3, *left*) as correctly and quickly as possible. When the subject had clearly constructed the 3D structure in the brain, he was to quickly press the mouse button. A white background was then presented for 150 ms, to remove any afterimage. The response time was limited to 30 s; if that time was surpassed with no response, the next trial was automatically started.

A 3D figure then appeared (Fig. 3, *right*). The subjects had to indicate whether the new figure corresponded to the mentally imaged 3D figure by clicking the mouse button: left = correct and right = incorrect. The time to indicate correct or incorrect (each 50%, randomized order) was limited to 10 s. If the time was surpassed, the next trial was automatically started.

Subjects were instructed to press the button as quickly and accurately as possible, and to concentrate on the next trial, regardless of their past mistakes. Practice trials (at least 10 trials) were performed to familiarize subjects with the task procedure; the actual task was then performed. A block was 54 trials; a total of two blocks of 108 trials were performed. The dummy trials (initial 5 trials) were added to the first block. Subjects had a rest period of almost five minutes between the blocks.

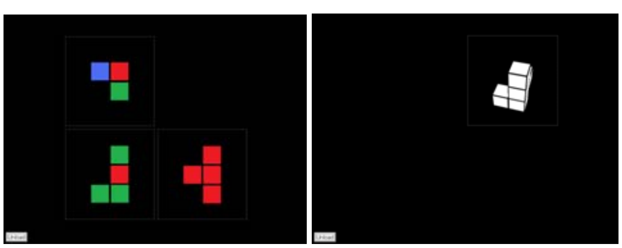

Figure 3. The figures used in task 4). Typical example of 2D (*left*) and 3D (*right*) figures. After a button was pressed, the 3D figure appeared.

Finally, the subjects completed a questionnaire to indicate their subjective feeling with regard to the effect of the three conditions (control, shadow, and color) on reaction time.

5) *3D image from non-colored 2D figures*

The procedure for this task was the same as that for task 4) above. The 2D figures shown were white. The angles of the 3D figures in the matching phase had 25 variants: 15, 30, 45, 60, and 75 degrees in both azimuth and elevation (one block = 50 trials; total three blocks = 150 trials, randomized order of the angles). The dummy trials (initial 5 trials) were added to the first block.

E. *Data analysis*

Reaction time and selected answers during all tests were recorded for later analysis. Average reaction time and the percentage of correct responses were calculated for each subject. All reaction times in each task were averaged for each subject; the mean value was subtracted from each reaction time. The ground averages among all subjects were then calculated. All data were presented as mean ± SD. The instances of no response and the initial five trials of all tasks were excluded to prevent dispersion during the initial trials.

To evaluate the data of a repeated-measures design in subjects receiving all conditions, non-parametric Friedman tests [7] were performed. After the Friedman tests, Tukey tests were applied for multiple comparison. In tasks 1) and 2), Wilcoxon signed-rank test were performed. Statistical significance was assigned to differences producing $p < 0.05$.

III. RESULTS

A. *Optimal 3D angle*

Figure 4A shows the mean values of optimal view angles in a 3D object. The elevation (Fig. 4A, *left*) was significantly smaller than the azimuth (Fig. 4A, *right*). The elevation and the azimuth were 27.2 ± 2.7 and 44.8 ± 4.1 degrees respectively; the difference between the two angles was almost 17 degrees ($p < 0.01$).

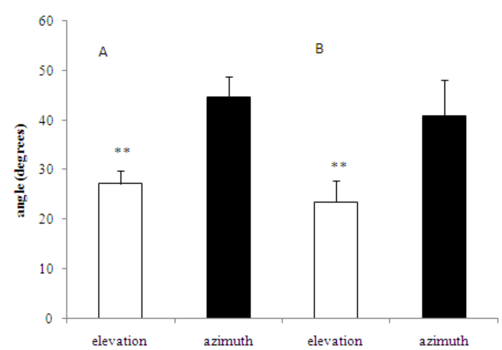

Figure 4. Mean angles of elevation and azimuth in 3D objects among all subjects. A: Optimal 3D angles. B: Angles of 3D image. ** shows $p < 0.01$.

B. Angle of 3D image

Figure 4B shows the mean values of view angles in a 3D image constructed from 2D figures. The elevation (Fig. 4B, *left*) was significantly smaller than the azimuth (Fig. 4B, *right*). The elevation and the azimuth were 23.6 ± 4.3 and 41.0 ± 7.2 degrees respectively; the difference between the two angles was almost 17 degrees ($p < 0.01$).

C. Simple reaction task

Simple reaction times to squares were evaluated. White, light gray, and dark gray squares in this task were consistent with those for the shadow condition in task 4); red, green, and blue squares corresponded with those for the color condition. The mean reaction time to the dark gray square was the slowest (262.4 ± 32.7 ms). The reaction time to red was longer than those to green and blue (254.0 ± 25.9 vs. 246.1 ± 16.4 and 246.8 ± 12.4 ms). However, there were no significant differences among all squares in this experiment.

D. 3D image from colored 2D figures

The mean percentage for the analyzed data was 99.2 % of all trials in the brain-image phase. Those data were used for the following data analyses. In the 3D-matching phase, the mean correct answer was 83.6 % of all trials, and the result of a Friedman test among the three conditions was not significant. All response data were used for the following data analyses.

1) Mental image phase

Figure 5 shows the mean reaction time among all subjects during the first 3D-image phase. The angles (30, 45, or 60 degrees) of the later 3D matching phase were not related to this phase, because the 2D figures alone appeared without a 3D figure. Actually, the reaction times among those angles in each condition were almost the same. Therefore, all data of those angles in each condition were used for the following statistics. The mean reaction times of the raw data were 11.0 ± 3.2, 9.3 ± 2.2, and 9.0 ± 2.2 seconds in the control, shadow, and color conditions, respectively. The result of the Friedman test for the reaction time among three conditions was significant ($\chi^2 = 13.6$, $p < 0.01$). Post hoc analysis showed the reaction times in the shadow ($p < 0.05$) and color ($p < 0.01$) conditions were significantly shorter than those in the control condition. The differences in reaction times between the shadow and color conditions were not significant.

2) Phase of matching of 3D objects

The results of the Friedman test and post hoc analysis on reaction time in the 3D-matching phase were not significant among the angles (30, 60, and 90 degrees) used in the three conditions (i.e., 9 combinations).

3) Questionnaire

The scores on the questionnaire were 3.0 ± 1.4 in the control condition, 4.9 ± 1.0 in the shadow condition, 5.8 ± 1.4 in the color condition. The result of the Friedman test on the scores among the three conditions was significant ($\chi^2 = 7.5$, $p < 0.05$); post hoc analysis showed the score in the color condition was significant, compared with that in the control condition ($p < 0.05$).

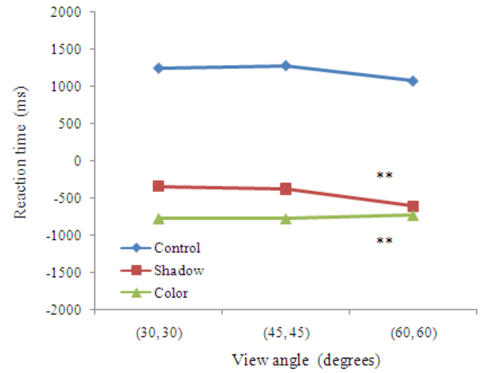

Figure 5. Mean reaction time in constructing a 3D image from 2D figures among all subjects. ** ($p < 0.01$) vs. control condition.

E. 3D image from the non-colored 2D figures

The mean percentage for the analyzed data was 98.9 % of all trials in the mental image phase, and 98.3 and 98.1 % for responses of "correct" and "incorrect" respectively in the 3D-matching phase. The mean correct answer was 81.4 % of all trials; the Friedman test result among the combinations of angles was not significant. All response data were used for the following data analyses.

1) Mental image phase

In the reaction time of the 3D image phase, the Friedman test and post hoc analysis results were not significant for all combinations of angles in elevation and azimuth. The angles of the later 3D matching phase were not related to this phase, because the 2D figures appeared without a 3D figure. Therefore, the insignificant difference among angles of this phase shows the task equality in the figures used.

2) Phase of matching of 3D objects

Figure 6 shows the mean reaction time among all subjects on various viewpoints, in the matching phase between the imaged and visual 3D objects. The mean values of the optimal angle and 3D image in tasks 1) and 2) were also superimposed in Fig. 6. The mean reaction times of the raw data were 1.4 ± 0.8 and 2.1 ± 1.2 seconds in the correct and incorrect 3D objects, respectively.

In the correct 3D objects (50 % of all tasks, Fig. 6A), the Friedman test result for the reaction time was significant among the combinations of angles ($\chi^2 = 42.2$, $p < 0.05$). The reaction times were especially shortened for (15, 15), (30, 30) and (45, 30) degrees in the elevation and azimuth, and the bilateral angles, showing the border line at 45 degrees in the azimuth. Post hoc analysis showed the reaction time in the combination of (15, 15) degrees in the elevation and azimuth was significantly shortened, compared to that of (75, 60) degrees ($p < 0.05$).

In the incorrect 3D objects (50 %, mirrored image, Fig. 6B), the results of the Friedman test and post hoc analysis for reaction time were not significant in the combinations of view angles. The view angles with shortened reaction times were varied. However, the reaction time had the tendency to decrease on the viewpoints of 15 degrees in the elevation and (60, 60) degrees in the elevation and azimuth.

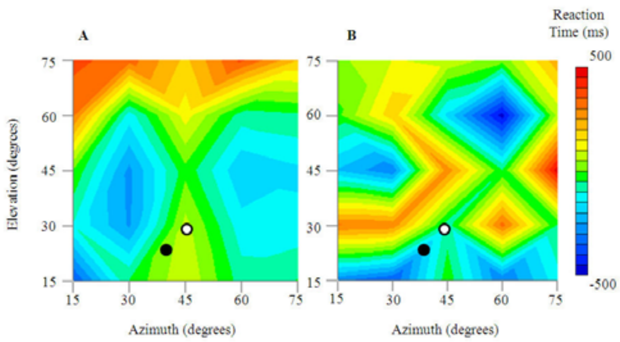

Figure 6. Contour maps of the reaction time of the 3D matching phase. The horizontal line shows azimuth, and the vertical line shows elevation of the view angles. A: correct 3D objects; B: incorrect 3D objects. Open circle: the optimal 3D angle in task 1); closed circle: the 3D image angle in task 2).

IV. DISCUSSION

The cognitive process in the brain during 3D tasks was investigated. The features of this experiment were specially designed to evaluate the ability of subjects to construct a 3D image from 2D figures and to match the actual 3D object and the imaged one. The results show several points (e.g. optimal presentation angles and the effects of depth cues) which could facilitate the cognitive process in the 3D human-computer interface.

The cognitive response in various angles of 3D. Using reaction time, the optimal angles for a 3D presentation were explored in task 5) (Fig. 6). In the correct 3D objects (Fig. 6A), the reaction times in the combinations around 15 to 30 degrees in elevation and azimuth were significantly shortened. In addition, on the bilateral angles bordered at 45 degrees in azimuth (symmetrical), the reaction times were shortened, although the reaction time was significantly increased when the depth angle was large (i.e. 75 degrees in elevation). If a simple rotation based on a 3D mental image is performed, the reaction time is linearly increased as the angles of the 3D object increase, as in the previous study [3]. Therefore, other functions in the brain activities must be activated in this task; when mentally constructing the 3D image, the subject might have been mainly referencing the side as well as front views. This result indicates a top-down process after or during the process of understanding the 3D structure; there might have been a bottom-up process from the visual stimuli [8] as well. The slow response at 45 degrees might be also due to the difficulty in instantaneously discriminating the 3D figures because of the visualization of symmetrical 3D blocks.

Relationship between short reaction time and the optimal or imaged 3D angles. The angles of imaged 3D objects in task 2) were almost consistent with the optimal ones in task 1); the imaged 3D angles might be slightly smaller. This might mean the imaged angles were mainly based on the frontal view of 3D. In addition, the difference of elevation and azimuth was almost 17 degrees, indicating that a significant (but not large) difference in view angles is preferable in the tasks 1) and 2).

Whereas the author hypothesized that the imaged or optimal 3D angles in tasks 1) and 2) would be similar to those with the shortest reaction time in task 5), the values were not always the same. Therefore, in addition to the imaged (or optimal) 3D angles, other abilities must have been operating in this experiment. The likely possibility is a top-down process [8] based on the front or side (not top) view of the imaged 3D from 2D figures in task 5), as described above. The time limitation for task 5) might have caused the difference in results, because it required a quick response.

Effects of depth cues. During the process of constructing 3D images from 2D figures, the reaction time was significantly shorter in the color condition than in the control condition (Fig. 5). It is generally known that depth cues help to understand 3D shapes [1]. In this experiment, whereas the depth cues added to the 2D figures were not directly shown in the visual 3D structure, even in the construction of the 3D image in the brain, the depth cues in the 2D figures facilitated the cognitive process. In addition, the scores on the subjective questionnaire in task 4) correspond to the results of reaction times in the three conditions.

On the other hand, once the 3D image was constructed in the brain, no significant effects of the depth cues were observed during 3D matching phase in task 4). Therefore, the constructed 3D image or the mental understanding of 3D structures (not only based on the visual stimulus in a bottom-up process) might not be significantly related to the later tasks, although the effects of the depth cues might appear in the more complicated tasks.

Brain activity. The cognitive process in constructing a mental image would be based on the ability of the visual cortex to sense the stimulus [9]. The activation of the left parietal cortex was demonstrated during the mental rotation task [4]. Furthermore, this experiment might produce brain activity in the frontal lobe and the parietal area related to spatial working memory [10, 11] as a result of the subjects' maintaining, recalling, and operating 2D figures. Searching for the same 3D structure as the newly presented one and judging whether it is correct or false requires high brain processes. A mentally constructed 3D figure without direct visual information might require further working memory.

Perspectives. There are some cases when one must estimate and judge a 3D shape from multiple plane pictures from MRI or CT scans in a clinical assessment. In such cases, an accurate cognitive process is required in real time, and cognitive process abilities may be decreased when one is fatigued, for instance, when performing brain surgery. The results of this experiment might indicate an effective method of presentation for quick recognition in plane MRI or CT scans, or the optimal angles of a 3D structure in such a case.

As a limitation, simple 3D figures based on five cubes were used for this experiment. The view angles were also limited in this experiment. When more complicated 3D structures on wider view angles are used, greater differences might be caused in the reaction time and accuracy for the 3D tasks.

In conclusion, this experiment was performed in order to evaluate specific brain activities using human responses in constructing 3D images from plane materials, with comparison to a general mental rotation task [3]. The results demonstrated several significant points with respect to optimal presentation angles and the effects of depth cues with shadow and color in facilitating the cognitive process in human-computer interfaces. In future studies, further investigation into effective methods for 2D and 3D presentations, from the viewpoint of human-computer interface and based on the evaluation of human cognitive function, will be required.

ACKNKOWLEDGEMENT

This study was funded by a Grant-in-Aid for Young Scientists (B) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (KAKENHI, 20700392).

REFERENCES

[1] S. Grossberg, "Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures," Psychol Rev, vol. 104 (3), pp. 618-658, 1997.

[2] K. Kashihara, "Evaluation of the cognitive process during mental imaging of two- or three- dimensional figures," the proceedings of the second international conferences on advances in computer-interactions, pp. 126-129, 2009.

[3] R.N. Shepard and J. Metzler, "Mental rotation of three-dimensional objects," Science, vol. 171 (972): pp. 701-703, 1971.

[4] M.S. Cohen, S.M. Kosslyn, H.C. Breiter, G.J. DiGirolamo, W.L. Thompson, A.K. Anderson, S.Y. Brookheimer, B.R. Rosen, and J.W. Belliveau, "Changes in cortical activity during mental rotation. A mapping study using functional MRI," Brain. vol. 119, pp. 89-100, 1996.

[5] T. Vetter, T. Poggio, and H.H. Bulthoff, "The importance of symmetry and virtual views in three-dimensional object recognition," Curr Biol, vol. 4, pp. 18-23, 1994.

[6] I. Biederman and P.C. Gerhardstein, "Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. J Exp Psychol Hum Percept Perform," vol. 19, pp. 1162-1182, 1993.

[7] S.A. Glantz, Primer of Biostatistics (4th ed.), New York: McGraw Hill, 1997.

[8] G. Ganis, H.E. Schendan, and S.M. Kosslyn, "Neuroimaging evidence for object model verification theory: Role of prefrontal control in visual object categorization," Neuroimage, vol. 34, pp. 384-98, 2007.

[9] S.M. Kosslyn, "Mental images and the brain," Cognitive Neuropsychology, vol. 22, pp. 333–347, 2005.

[10] E.E. Smith and J. Jonides, "Storage and executive processes in the frontal lobes," Science, vol. 283 (5408), pp. 1657-1661, 1999.

[11] V.A. Diwadkar, P.A. Carpenter, and M.A. Just, "Collaborative activity between parietal and dorso-lateral prefrontal cortex in dynamic spatial working memory revealed by fMRI," Neuroimage, vol. 12, pp. 85-99, 2000.