

Research Project for Human-centered Utilization of Visual Information for Surrounding Computing (Report for 2007 academic year)

Keizo SHINOMORI*, Mamoru OKADA**, Yoshimasa KIMURA†
and Xiangshi REN***

Department of Information Systems Engineering, Faculty of Engineering,
Kochi University of Technology

185 Miyanokuchi, Tosayamada-town, Kami-city, Kochi 782-8502, JAPAN

E-mail: * shinomori.keizo@kochi-tech.ac.jp, ** okada.mamoru@kochi-tech.ac.jp,
*** ren.xiangshi@kochi-tech.ac.jp

Abstract: Our group conducts collaborative research projects in order to enhance the human-centered utilization of visual information, especially for Surrounding Computing (SC). We perform our investigation in three research areas; (1) Applicable human perception area, (2) Applicable human operation area and (3) Image recognition algorithm area. In 2007 academic year, we concentrated four research themes.

The first theme is about temporal feature of S-cone pathway described by Impulse Response Function (IRF) in the applicable human perception area. The impulse response function (IRF) is a temporal response of entire visual system to an extremely short flash, measured by a psychophysical method. In this year, it was especially focused on IRFs of S-cone pathways, in which there are two different pathways for S-cone ON-response and S-cone OFF-response. As the results, the S-cone ON-IRF was slower than the luminous IRF but faster than the OFF-IRF. The difference between S-cone ON- and OFF-IRFs were statistically significant.

The second theme is about the effect of subjective biases in steering tasks evaluation in the applicable human operation area. Current studies on steering tasks focus on the effect of system factors (i.e., path width and amplitude) on the movement time and steering law's related applications. We conducted a series of experiments to further explore the effect of different operational biases (bias speed or accuracy) on steering completion time and standard deviation for two steering trajectory shapes, i.e., a straight steering task and a circular steering task, and then, establish a new model accommodating system and subjective factor in steering tasks. Empirical results showed that the new model is more predictive and robust than the traditional steering law.

The third theme is about sign language. The sign language is widely used as a communications means of the hearing impaired. In this year, we focused the research to recognize yubimoji by the distance evaluation method. Yubimoji is a system of manual kana used as part of Japanese Sign Language. Manual size and the speed of expressing a word by yubimoji are different by a person. Therefore, the recognition method that is robust to be influenced by them is needed. The movement of expressing yubimoji is followed by using two dimension measurement software, and this research was aimed for inspecting about recognition method of yubimoji by its data.

† Currently, the third author belongs to: Department of Computer System Technology, Faculty of Computer and Information Sciences, Sojo University, and his e-mail address is: kimura@cis.sojo-u.ac.jp.

The fourth theme is about the technology to retrieve similar images to the presented query image as similar image retrieval. We propose the image navigation that repeats narrowing the image over the multistage on the multidimensional psychology space, which reflects subjectivity evaluation value of the non-similar degree between images. The result shows the effect of this technology and shows whether the axis of the multidimensional psychology space refers any feature on the multidimensional psychology.

1. Structure of the project

1.1. Purpose

Visual information plays the most important role as a source of information processing related to humans and is mostly used form to present information in information systems concerning humans in any process. Thus, in the most of information systems concerning human, the information between them and human are mainly presented in visual information. Especially, color is extensively used in information systems and technologies. However, it has not been well investigated about the relationship between the utilization of visual information in such systems and human visual information processing and/or human behavior, neither in scientific aspects nor in applicable aspects. Especially for Surrounding Computing (SC), it is required to have all-around knowledge and applicable algorithms about perception to shape and color, visual processing with behaviors, automatic image-and-movie processing and human perception and recognition.

Our group conducts collaborative research projects in order to enhance the human-centered utilization of visual information, especially for Surrounding Computing, in collaboration with the Graduate School of Department of Information Systems Engineering in KUT. In this three-years research project, we perform our investigation in three research areas; (1) Applicable human perception area, (2) Applicable human operation area and (3) Image recognition algorithm area. In those research areas, we especially focus on those sub-projects; (1) Prediction of human response speed for visual stimulation (2) Automatic perception system for human face expression recognition, (3) Prediction system for chromatic area effect, (4) Prediction system for human pointing performance on Human-Computer-Interaction (HCI), (5) Automatic motion retrieval system for the movement of a sign

language translator, and (6) Image search system for similar pictures.

From the systematic view to results of these sub-projects, we try to enhance easy and accurate utilizations of visual information and construction of comfortable human interface. Because this research process also includes some of important tasks those should be investigated, the knowledge and information obtained in this project must be useful to apply to other problems in computer applications. We also expect that Ph.D. students and masters students in the graduate school will join our research projects and will acquire high ability as specialists in color science and engineering.

1.2. Outline in 2007 academic year

The main theme of this project continues in three years and 2007 academic year is the first year. In this year, we performed these sub-projects in one of three research areas below;

(1) Applicable human perception area:

Temporal feature of S-cone pathway described by Impulse Response Function. (Leader: Keizo SHINOMORI)

Human Face Perception and Face Expression Perception. (Leader: Keizo SHINOMORI)

(2) Applicable human operation area:

Investigation of the Effect of Subjective Biases in Steering Tasks Evaluation. (Leader: Xiangshi REN)

(3) Image recognition algorithm area:

Yubimoji recognition by distance evaluation. (Leader: Mamoru OKADA)

Hand shape recognition by an outermost point method. (Leader: Mamoru OKADA)

Similar Image Retrieval Using Multidimensional

Psychological Space. (Leader: Yoshimasa KIMURA)

Through collaborative work in these three core areas, we expect to establish an easy and accurate means of utilization of visual information for Surrounding Computing.

We have made progresses in all themes. Three papers have been accepted and will be published soon in journals. Some results have been presented in international meetings as shown in the next section.

1.3. Achievement in 2007 academic year

The 2007 academic year is the third year in our joint project, thus we have some final achievements directly concerning to the project.

Journal paper and book;

- J1. Shinomori, K. and Werner, J.S. (2008). The impulse response of S-cone pathways in detection of increments and decrements, *Visual Neuroscience* Vol. 25, No.3, pp.341-347.
- J2. Ren, X., Zhou, X. (2008). An Investigation of Subjective Operational Biases in Steering Tasks Evaluation, *Behaviour & Information Technology*, Taylor & Francis, in printing.
- J3. Ren, X., Zhang, X., and Kyo, K (2008). Quantifying the Learning Effect in Human Performance Models, *International Journal of Innovative Computing, Information and Control (IJICIC)*, vol.4, No.9, pp2419-2429.

International conference publishing reviewed proceedings;

(1) Applicable human perception area:

1. Song, M. and Shinomori, K. (2007). Does the attractiveness-bias exist in the smile expression perception?, *IPSI SIG Technical Report*, 2007-CVIM-160(5), pp.23-28.
2. Shinomori, K. and Werner, J.S. (2007). The Impulse Response Functions for S-Cone Increments and Decrements, *Proceedings of NEINE'07 (the International Conference on Next Era Information Networking, Shanghai, China, 23-24 September 2007)*, pp.292-297.
3. Song, M., Watanabe, T. and Shinomori, K. (2007).

Attractive Person Presents Smile with Lower Physical Intensity Estimated by Graphic Morphing Technology, *Proceedings of NEINE'07*, pp.307-312.

4. Shinomori, K. (2007). Aging effect to inter-generational and intra-generational variations in visual sensitivities to luminous and chromatic (S-cone) flashes, *Proceedings of the International Conference on Kansei Engineering and Emotion Research (KEER 2007)*, A-19.
5. Song, M. and Shinomori, K. (2007). Affects of Impression of Face to Smile or Angry Perception, *Proceedings KEER 2007*, A-22.
6. Shinomori, K. (2008). Temporal feature of S-cone pathway described by Impulse Response Function, *Vision (the Journal of Vision Society of Japan)*, Vol.20, No.2, pp.67-71.

(2) Applicable human operation area:

7. Ren, X., Zhou, X. and Liu, Z. (2007). An Empirical Evaluation of Seven Mice for Scrolling Tasks, *Proceedings of the 2007 IEEE International Conference on Mechatronics and Automation (ICMA 2007)*, pp.582-586.
8. Kong, J. and Ren, X. (2007). Information Processing Rate in Human-Computer Interaction, *Proceedings of The Second International Conference on Innovative Computing, Information and Control (ICICIC2007) in CD-ROM*, IEEE computer Society.
9. Fukutoku, F., Ren, X., and Zhou, X. (2007). An Empirical Evaluation of Upper Bound Limit of Width for Steering Task, *Proceedings of NEINE'07*, pp.196-198.
10. Zhou, X., and Ren, X. (2007). An Investigation of Subjective Operational Biases in Steering Tasks Evaluation. *Proceedings of SJCIEE2007*, p.341.
11. Fukutoku, F., Ren, X., and Yin, J. (2007). The Upper Limit Size of Path Width for Steering Law. *Proceedings of SJCIEE2007*, p.339.
12. Higaki, T., Ren, X., and Zhou, X. 2007. An Investigation of Influence of Different Start Position for Steering Tasks. *Proceedings of SJCIEE2007*, p.340.

(3) Image recognition algorithm area:

13. Tanahashi, M., Hiromitsu, D. and Okada, M.

- (2007). Hand shape recognition by an outermost point method, *Proceedings of NEINE'07*, pp.176-177.
14. Tanaka, T., Nagano, Y. and Okada, M. (2007). Yubimoji recognition by distance evaluation, *Proceedings of NEINE'07*, pp.180-181.
 15. Suemitsu, M., Hirai, T. and Okada, M. (2007). Characteristic extraction of the dace expression by optical flow, *Proceedings of NEINE'07*, pp.174-175.
 16. Hisashi, S., Fujitani, K. and Okada, M. (2007). Presumption of frame which feelings are expressed most through expression change of a face, *Proceedings of NEINE'07*, pp.178-179.
 17. Hirose, K., Ohnishi, N. and Kimura, Y. (2007). Presumption of frame which feelings are expressed most through expression change of a face, *Proceedings of NEINE'07*, pp.465-468.

In following chapters, we introduce important achievements in original shape with the authorization of publishers. In the applicable human perception area, No. 6 article in the list of international conference proceedings shown above is presented. In the applicable human operation area, No. 2 paper in the list of the journal paper and book is presented. In image recognition algorithm area, No. 14 article in the list of international conference proceedings is presented. Additionally, in this area, summary of the research about similar image retrieval will be presented.

2. Research achievement in applicable human perception area

We summarize the research achievement in applicable human perception by one paper that was published by Keizo Shinomori as the journal paper for *Vision (the Journal of Vision Society of Japan)*(ref.6 in 1.3). The paper is shown below.

Temporal feature of S-cone pathway described by Impulse Response Function. **Keizo SHINOMORI**

2.1 Introduction

The impulse response function (IRF) is a temporal response of entire visual system to an extremely short

flash, measured by a psychophysical method. In experiments with control of luminous and chromatic changes of the flash, it is possible to obtain the IRFs mainly determined by the achromatic or chromatic pathway.

I especially focused on IRFs of S-cone pathway. Because S-cone pathway is anatomically and physiologically different from achromatic pathway and chromatic pathways originated from L and M cones. Especially, in S-cone pathway, there are different two pathways for S-cone ON-response and S-cone OFF-response. In S-cone ON-response, excitatory (ON) signal from S-cones with inhibitory (OFF) L+M cones signal are going through Small bistratified ganglion cells and connected to Konio-cellular pathway in LGN¹⁾. On the contrary, in S-cone OFF-response, excitatory signal from S-cones with inhibitory L+M cones signal are going through S-cone ON bipolar cell but they enter to a sign inversion connection of Monostatified ganglion cells those are a few and large sparse cells²⁾.

By these differences in S-cone ON- pathway and OFF-pathway, it is expected that impulse responses of S-cone are different with the responses of L-cone and M-cone as previously expected³⁾. Also, the responses of S-cone ON and S-cone OFF can be different each other. Thus, I measured the chromatic impulse response function of an isolated human S-cone pathway in double pulse method. About chromatic impulse response functions, mostly impulse responses were measured with red and green isoluminant pulses.^{4,5)} Chromatic (blue) impulse response functions have also been reported, but not with S-cone isolating stimuli⁶⁾. Here, I measured the chromatic impulse response functions of an isolated human S-cone pathway. Additionally, I compared impulse response functions of putative ON- and OFF-pathways using chromatic increments and decrements.

As the results, the S-cone ON-IRF was slower than the luminous IRF but faster than the OFF-IRF. The difference between S-cone ON- and OFF-IRFs were statistically significant.

2.2 Methods

2.2.1 Individual luminance and individual tritan line measured on CRT

Impulse response functions were derived from thresholds for a series of double-pulses in which the pulses were chromatically modulated on individual tritan lines at constant luminance.

First, I measured an individual luminance by obtaining minimum flicker points between phosphors as our previous works^{7,8)}. Luminance of one phosphor (ex. blue phosphor) of the CRT was fixed at 6.9 cd/m^2 and presented stimulus consisting of the light only from that phosphor and the light only from the other phosphor (ex. red phosphor) in chromatic alternation flicker (18Hz) of a temporal square wave. The stimulus was a 0.64-degree to 2.77-degree annulus surrounding a central fixation cross. Because ocular media transmission of the red phosphor is affected negligibly by lenticular senescence, and because the optical system produced an image of the stimulus in the plane of the pupil that was smaller (2.5 mm diameter) than that expected for observers spanning our age range⁹⁾, I expected that retinal illuminance was equated across observers.

In second, I measured individual tritan lines for each observer by color matching method with a strong adaptation to S-cones by the 420 nm light presented in a Maxwellian view. The observer was asked to find a match between two rectangular patches by adjusting the angle of the line around the white point and the intensity of one of the two test patches. All observers reported that they reached to the metameric match, which meant the match both in color appearance and brightness.

2.2.2 Apparatus and stimuli

Two-pulse thresholds were measured for stimuli that were modulated in chromaticity along a tritan line, changing from the white background (CIE (x, y) = (0.33,0.33)) in 10 cd/m^2 ($1.69 \log \text{ Td}$) toward the short-wave spectrum locus (blue pulse) or the long-wave spectrum locus (yellow pulse). The two pulses were 6.7 ms each for blue pulses or 40 ms each for yellow pulses and separated by interstimulus intervals (ISI) ranging from 20 - 360 ms. A central fixation cross defined four positions of the test stimuli

which were located 1.70 deg to one side or the other and 1.70 deg above or below the center of the fixation cross. The test stimulus was a Gaussian patch, 2.26 deg diameter at 1 SD, chosen as the test spatial profile to eliminate artifacts caused by spatial transients.

These stimuli were presented on a CRT display (Sony GDM-200 PS) operating at a 150 Hz frame rate that was controlled by a video board with 15-bit resolution (Cambridge Research Systems, VSG 2/4). An aperture was placed before the telescope objective and the focused image of the CRT and adapting field were 1.5 mm in the plane of the eye pupil. Observer position was maintained with a dental-impression bite-bar.

2.2.3 Procedure

I tested four normal observers (3 males and 1 female), ranging in age from 21.3 to 40.1 years.

Each session began after 5 min dark adaptation and 5 min adaptation to a 10-cd/m^2 equal-energy white background. Two pulses were presented on the screen, preceded by a high-pitched tone and followed by a low-pitched tone. The observer's task was to indicate in which of four quadrants the stimulus was detected by pressing one of four correspondingly arranged buttons. The stimulus was a chromaticity change in one Gaussian patch from equal-energy white along the individually determined tritan line.

In S-cone ON impulse response function measurement, only two blue pulses were presented in one session. The change was a double pulse (6.7 ms with interstimulus intervals from 20 to 360 ms) in which the two flashes were modulated equally in chromaticity at constant luminance toward blue, resulting in increased S-cone excitation. This 4-alternative forced-choice task was combined with a two-down, one-up staircase in which staircases for each ISI were interleaved. Thresholds for each ISI were based on the last four of six reversals corresponding to a 70.7% probability of detection. This was repeated in at least 4-6 sessions per observer.

In the case of S-cone OFF impulse response function measurement, the procedure was basically the same except only two yellow pulses were presented in one session, and 6 frames instead of one frame were used for each yellow pulse, because of insufficient

chromatic change to yellow in a single frame.

The double pulse method is the method to derive a shape of an IRF with some hypothesis by measuring the change of thresholds of flashes when an inter-stimulus interval of two flashes is changed. From the threshold change as a function of inter-stimulus interval, the shape of the IRF can be estimated precisely in some degree. I employed the model of the IRF by Burr and Morrone⁵⁾ as described in our previous works^{7,8)}.

2.3 Results

Figure 1 shows results of threshold as a function of ISI measurement by S-cone ON flashes (top panels) and S-cone OFF flashes (bottom panels). Not like luminous IRFs, both S-cone IRFs are mono-phasic and the durations protracted compared to the bi-phasic and tri-phasic IRFs that characterize response to an achromatic double pulse⁷⁾. On this observer, the time to the peak are 64 ms for S-cone increment IRF and 145 ms for S-cone decrement IRF, compared with 21.3 ms with luminance modulation (shown in Ref. 7, Fig.7). With luminance modulation, the first excitatory phase is followed by an inhibitory phase making the excitatory duration easy to define. For the S-cone IRF, the duration of the excitatory phase will be operationally defined as the value on the descending slope corresponding to 5% of the peak amplitude.

In the experiment, it was not possible to obtain sufficient S-cone modulation in a double-pulse defined by single frames because of the luminance limit of the CRT phosphors. However, because the S-cone IRF is slow, it was possible to use multiple frames (typically 6 frames) to define each pulse without affecting the shape of the IRF. In this case, I defined each pulse as one series of successive frames of the IRF calculated as the summation of S-cone IRFs with time delay by each frame (6.67 ms frame rate). In Figure 1, this IRF is shown by the smaller curve in which each pulse was defined by six frames for S-cone decrement IRF measurement. This method of measurement and

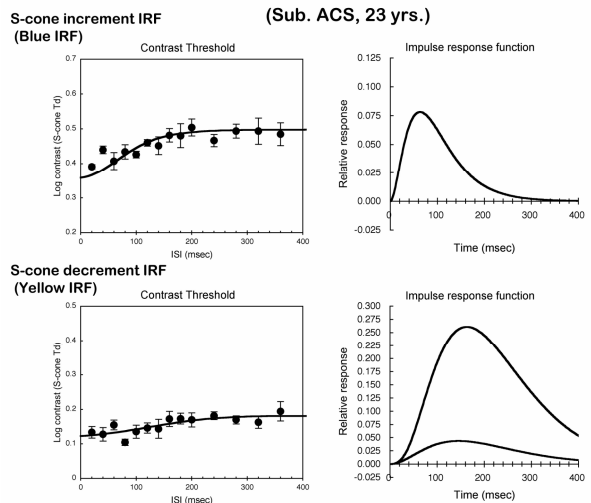


FIGURE 1: Threshold data with the model fit and impulse response functions (IRFs) calculated by the model for one observer (Observer ACS, 23 years old). Top panels are for S-cone increment IRF and bottom panels are for S-cone decrement IRF, respectively. In the case of S-cone decrement IRF, because one pulse consisted of 6 frames, linear summation of 6 response functions with 6.67 ms delay each was used to calculate the expected threshold curve from the shape of IRF (denote by black curve in the threshold data panel).

calculation assuming linear summation is justified because the double-pulse method itself has to employ this assumption to obtain the IRFs.

2.4 Discussion

In order to confirm the difference between S-cone increment and decrement IRFs, two characteristics of IRFs were compared on each observer in a statistical significance. Figure 2 and 3 show the difference between S-cone increment and decrement IRFs in terms of relative peak amplitude and peak time, respectively. Differences in all characteristics are statistically significant in 5 % significance level when the data were paired on each observer.

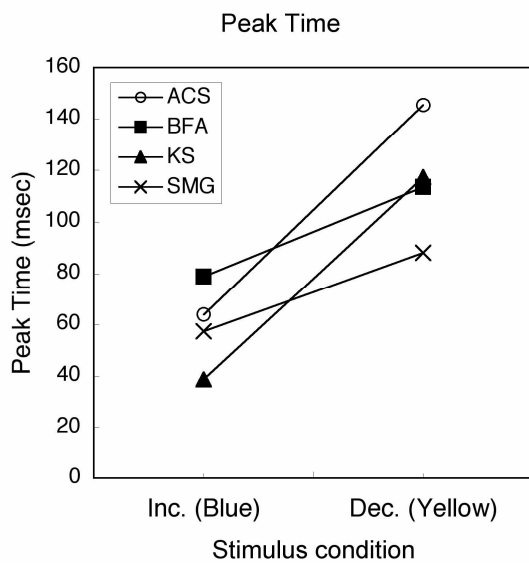


FIGURE 2: Relative peak amplitude in S-cone increment and decrement IRFs compared on each observer.

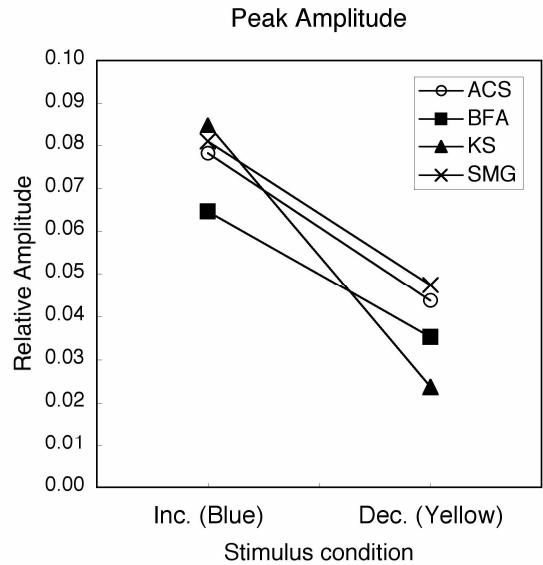


FIGURE 3 Peak time in S-cone increment and decrement IRFs compared on each observer.

There are, however, some individual differences between observers. For observer BFA, the difference is not so large, although for observer KS the difference is critical. I expect that this individual variation is caused by a usage of these S-cone increment and decrement pathways. In other words, these S-cone pathways are regularly not used to process visual stimulation concerning to temporal information. Thus, it is possible that these S-cone pathways are not well controlled in terms of temporal characteristics but only controlled in terms of chromatic characteristics. This hypothesis can be supported by the fact that S-cones are sparse and do not exist at the central fovea. Also, as described at the first part of this paper, S-cone pathways are quite different with L-cones and M-cones pathways.

Acknowledgement

Main part of the presentation and this summary has already been described in ref.10. This work is based on the collaborative work with Professor John S. Werner in University of California, Davis. This work was supported by the Special Subsidies in Subsidies for ordinary expenses of private schools from the Promotion and Mutual Aid Corporation for Private Schools of Japan to K. Shinomori and by a National Institute of Health grant (NIA AG04058) to J.S. Werner.

References

1. D.M. Dacey and B.B. Lee, The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type, *Nature*, Vol. 367, pp.731-735, 1994.
2. D.M. Dacey, B.B. Peterson and F.R. Robinson, Identification of an S-cone Opponent OFF Pathway in the Macaque Monkey Retina: Morphology, Physiology and Possible Circuitry, *Investigative Ophthalmology & Visual Science*, vol. 43, E-abstract 2983, 2002.
3. H.E. Smithson and J.D. Mollon, Is the S-opponent chromatic sub-system sluggish?, *Vision Research*, Vol. 44, pp.2919-2929, 2004.
4. K. Uchikawa and T. Yoshizawa, Temporal responses to chromatic and achromatic change inferred from temporal double-pulse integration, *Journal of Optical Society of America / A*, Vol. 10, pp.1697-1705, 1993.
5. D.C. Burr and M.C. Morrone, Impulse-response functions for chromatic and achromatic stimuli, *Journal of Optical Society of America / A*, Vol. 10, pp.1706-1713, 1993.
6. K. Uchikawa and M. Ikeda, Temporal integration of chromatic double pulses for detection of equal-luminance wavelength changes, *Journal of*

- Optical Society of America / A, Vol. 3, pp.2109-2115, 1986.
7. K. Shinomori and J.S. Werner, Senescence of the temporal impulse response to a luminous pulse, *Vision Research*, Vol. 43, pp.617-627, 2003.
 8. K. Shinomori and J.S. Werner, Impulse response of an S-cone pathway in the aging visual system, *Journal of Optical Society of America A*/Vol. 23, pp.1570-1577, 2006.
 9. I.E. Loewenfield, Pupillary changes related to age, In *Topics in neuro-ophthalmology*, eds. H. S. Thompson, R. Daroff, L. Frisen, J.S. Glaser and M. D. Sanders, (Baltimore, 1979)(Williams and Wilkens, U.S.A., 1979) pp.124-150.
 10. K. Shinomori and J.S. Werner, The impulse response of S-cone pathways in detection of increments and decrements, *Visual Neuroscience* / Vol. 25, No.3, 2008, in printing.

3. Research achievement in applicable human operation area

We summarize the research achievement in applicable operation area by one book chapter that will be published by X. Ren and X. Zhou for *Behaviour & Information Technology* (Taylor & Francis)(ref.J2 in 1.3). The paper is shown below.

Investigation of the Effect of Subjective Biases in Steering Tasks Evaluation. *Xiangshi REN and X. ZHOU*

Abstract

The Steering law is a model for trajectory- based tasks, such as drawing and writing in GUIs. Current studies on steering tasks focus on the effect of system factors (i.e., path width and amplitude) on the movement time and steering law's related applications. We conducted a series of experiments to further explore the effect of different operational biases (bias speed or accuracy) on steering completion time and standard deviation for two steering trajectory shapes, i.e., a straight steering task and a circular steering task, and then, establish a new model accommodating system and subjective factor in steering tasks. Empirical results

showed that the new model is more predictive and robust than the traditional steering law.

3.1 Introduction

An important research branch in Human Computer Interaction (HCI) community is to seek to develop formal models [1, 2 and 3] useful for predicting or describing human behavior in interactions with computer systems. Fitts' law [2] is a very powerful model for pointing task evaluation and has found many uses in HCI [4, 5 and 6]. Fitts' tasks follow certain speed-accuracy trade-off rules, i.e., the more accurate the task to be accomplished, the longer it takes and vice versa.

In 1997, Accot and Zhai developed a new model for trajectory-based tasks, called steering law [7], deduced from Fitts' law. Similarly, steering law also follows certain speed-accuracy tradeoff rules and has been used for a number of studies [8, 9, 10, 11 and 12]. But still, there exist some issues with the steering model. For example, current studies about the steering model pay little attention to the subjective factor, i.e., the subject's operational biases toward speed or accuracy. If the stroke is drawn as accurately as possible, steering completion time may increase. Conversely, if the stroke is drawn as fast as possible, steering completion time may decrease. So, the traditional steering law only involving system factors, i.e., tunnel amplitude and tunnel width, is not precise enough to model human performance accurately.

The objective of this paper is to explore the comprehensive effect of different operational human biases in trajectory-based tasks and attempt to establish a new steering model involving not only system factors but also subjective factors.

3.2 Related works

Literature about motor behavior models in Human-Computer Interaction can be found in Mackenzie's paper [13], which gave a good summary of models of human movement relevant to HCI. However, subjective operational biases were not involved in it.

Recent studies about the effect of subjective operational biases in computer interaction tasks include the work of Zhai, Kong and Ren [14]. In that paper, the

different operational biases of subjects towards speed or accuracy in Fitts' target acquisition tasks were systematically and completely discussed. A series of related experiments had been conducted to explore the relationships between target utilization, task specification and subjective operational biases. Experimental results showed that target utilization was not only affected by operational biases, but also by target width and distance. Moreover, the effect of width was more significant than the effect of distance. W_e^1 model [15] could partly compensate for the subjective layer's effect, but not completely. A complete model which can predict the relationships between the subjective and objective layers does not exist.

The above mentioned study about different operational biases is based on spatial variability, i.e. the normal distribution of end-points, which lacks theoretical and empirical foundations. Based on temporal constraint (distribution of movement time data), a new model, called SH-Model [16], was presented to include system and human effects in Fitts' target acquisition tasks. Empirical analysis showed that the SH-Model is stronger than the Shannon model² [17] and W_e model.

According to the definition of "effective target width" [15] in Fitts' tasks, Kulikov and MacKenzie, et al. introduced spatial variability into steering tasks for straight tunnels and established "effective tunnel width" [18] of steering motions. Empirical results manifested that the newly built model is stronger and more natural than traditional steering law. Effective tunnel width partly reflected what the subject actually did, but the different biases of operations were not been systematically and comprehensively discussed and varieties in the shape of the tunnel had not been considered in that paper, e.g. considered that the natural arc motion of the hand warranted that a curved tunnel or round tunnel be included.

¹ W_e model, i.e., "effective target width" model is: $MT=a+b\log_2(A/W_e+1)$, where $W_e=4.133SD$

² Shannon model is: $MT=a+b\log_2(A/W+1)$ For both models, a and b are empirically determined constants, A is the pointing distance, W is the target width and is MT the mean time of task completion. SD is the standard deviation of end-points distribution.

This paper further comprehensively investigates the effect of five different operational biases on two steering tasks, i.e., a straight steering task and a circular steering task and introduces standard deviation as subjective factor. Finally, a new steering model is attempted to establish reflecting not only system factors but also subjective factors.

3.3 Experiment

3.3.1 Task

Our experiment takes a straight tunnel and a circular tunnel as two steering tasks (see Figure 1).

The difficulty for steering through a straight tunnel (see Figure 1.a) is $ID_s=A/W$, where A is the length of the tunnel, and W its width. For a circular tunnel, the movement amplitude A is equal to the circle circumference $2\pi R$, where R is the circle radius, so the difficulty for steering through a circular tunnel (see Figure 1.b) is $ID_c=2\pi R/W$. Steering law that models the relationship between completion time MT and tasks difficulty ID can be expressed in the following form: $MT=a+b \times ID_s$ for a straight tunnel, and $MT=a+b \times ID_c$ for a circular tunnel.

3.3.2 Biases

The earliest study about the operational biases of subjects was the work of Fitts and Radford [19].

They systematically manipulated the operational biases of three subjects towards accuracy (A), neutrality (N), and speed (S), by means of monetary award and penalty at 1 cent per point. Fitts' thesis was that human information capacity in motor responses is relatively constant despite different experimental manipulations, so their paper did not focus on the effect of W_e correction.

In the work of Zhai, Kong and Ren [14], the different operational biases of subjects, towards speed or accuracy, were systematically and completely discussed in Fitts' target acquisition tasks; while the subject of our research is the effect of subjective operational biases in trajectory-based tasks.

We would like to comprehensively investigate the effect of five different operational biases on the above two steering tasks. They are extremely accurate (EA), accurate (A), neutral (N), fast (F) and extremely fast

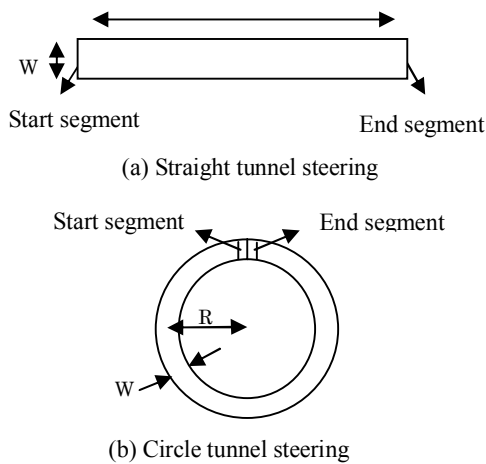


Fig. 1 Two steering tasks

(EF). The following verbal instructions corresponding to each operational bias are given by the experimenter to the participants: “Make a stroke along the tunnel as accurately as possible and do not worry about time or speed; try to avoid any error” in Condition EA; “as accurately as possible but keep some speed” in Condition A; “as accurately as possible and as fast as possible” in Condition N; “as fast as possible but keep some accuracy” in Condition F; and “as fast as possible and some errors are acceptable” in Condition EF.

3.3.3 Subjects

Ten subjects (7 male, 3 female, aged from 21 to 31) participated in the experiment. All participants had normal or corrected to normal sight. The participants performed the test using their preferred hand (all right handed).

3.3.4 Apparatus

The experiment was conducted on an IBM ThinkPad X41 Tablet PC with a stylus as the input device, running Windows XP. The screen size was 12.1 inches, with 1024×768 resolutions. Experimental software was developed with Java.

3.3.5 Design

The experiment was a 3×3 within-subjects factorial design. Group was a between-subjects factor with two levels (5 participants per group), with participants randomly assigned to Group 1 or Group 2. The within-subject factors were Task (linear vs. circular), Amplitude (250, 350, 450 pixels), Width (10, 25, 40

pixels). The direction of a linear task was towards right, and the direction of a circular task was clockwise.

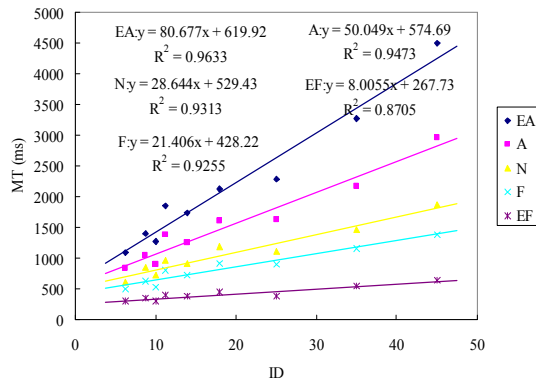
The participants in Group 1 first performed a linear task, while those in Group 2 did a circular task first. Each subject was instructed to repeat the experiment five times with different operational strategies, i.e., extremely accurate (EA), accurate (A), neutral (N), fast (F) and extremely fast (EF). The aforementioned verbal instructions corresponding to each operational bias were given by the experimenter to the participants before performing experiment.

The order of the EA, A, N, F, EF conditions was balanced by a Latin square pattern across each group of subjects. The order of the nine amplitude and width combinations was presented in random order to the participants in each operational bias. Each subject performed 9 strokes for each Amplitude/Width combination in each operational bias of the two tasks. So the total stroke number was 3 (tunnel amplitudes) × 3 (tunnel widths) × 9 (strokes) × 5 (operational biases) × 2 (tasks) × 10 (subjects) = 8,100.

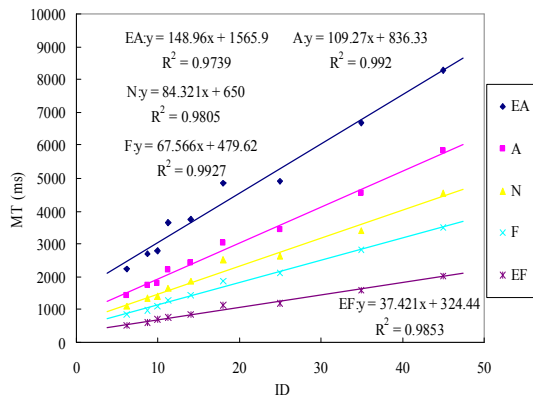
3.3.6 Procedure

The participants were first briefed on the purpose of the experiment. With the stylus as the input device, the subjects placed the Tablet PC on the desktop. Before the test, all subjects were allowed to perform some warm-up trials in each operational bias until they felt that they could begin the experiments.

Subjects performed two types of steering tasks: straight tunnel steering and circular tunnel steering (see Figure 1). At the beginning of each trial, the path to be steered was presented on the screen, in black. After placing the cursor to the left of the start segment and depressing the tip of the stylus, the subject began to draw a green line on the computer screen, showing the stylus trajectory. When the cursor crossed the start segment, left to right, the line turned blue, as a signal that the task had begun, the time was being recorded and the stylus trajectory was being sampled. When the cursor crossed the end segment, also left to right, the current tunnel disappeared and a new tunnel was presented to the subject. Lifting the pen tip up from the Tablet PC surface after crossing the start segment and before crossing the end segment would result in an invalid trial and that trial needed to be repeated. When



(a) Linear steering



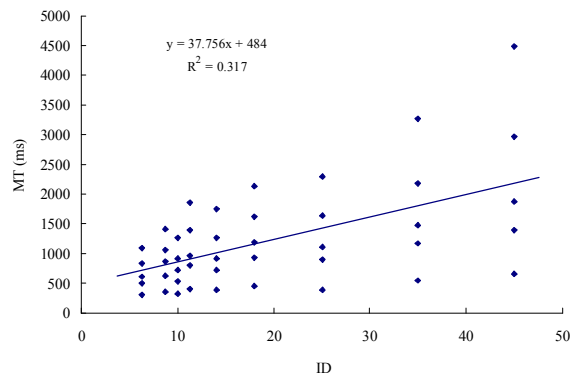
(b) Circular steering

Fig. 2 Mean completion time for each bias (EA, A, N, F and EF) as a function of difficulty in both linear and circular steering tasks

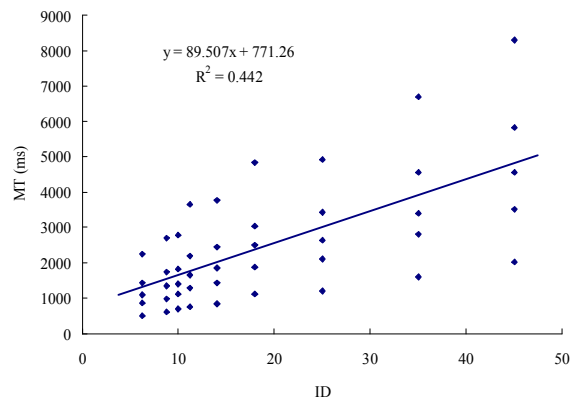
the cursor crossed the borders of the path, the line turned red, as a signal that the stylus trajectory was outside of the tunnel, but the current trial did not need to be redone.

3.3.7 Measurement

While the stroke was being made, the position of the cursor was sampled in intervals of 10 milliseconds. The dependent variables were: *MT* (time taken to move the cursor from the start line to the end line), *SD* (Standard Deviation: for the linear tunnel, *SD* is computed using the sampled y-values between the start line and the end line; for the circular tunnel, *SD* is computed using the distances between the sampled points and the center of the circular tunnel), and *OPM* (Out of Path Movement, percentage of sample points outside the tunnel border). For example, if 100 points were sampled and 10 of those points were outside the tunnel border, then *OPM* would be 10.



(a) Linear steering



(b) Circular steering

Fig. 3 Mean completion time for all biases combined as a function of difficulty in both linear and circular steering tasks.

3.4 Results

3.4.1 Movement time

ANOVA showed that a significant effect of bias ($F_{4, 405}=121.96$, $p<.00001$ for linear tasks, $F_{4, 405}=227.29$, $p<.00001$ for circular tasks) upon steering time. Mean steering time for EA, A, N, F and EF biases were respectively 2173.0, 1538.1, 1080.8, 840.3, and 421.8 milliseconds for linear steering tasks and 4433.4, 2939.8, 2273.2, 1780.3, and 1044.8 for circular steering tasks.

Further ANOVA analysis showed a significant interaction between index of difficulty and biases ($F_{32, 405}=4.42$, $p<.00001$ for linear tasks, $F_{32, 405}=4.98$, $p<.00001$ for circular tasks) (see Figure 2).

For linear steering, there was an interesting tendency that the more risky (faster paced) the operational condition was, the weaker the correlation between *MT* vs. *ID* was. That is to say, the R^2 values between *MT* vs. *ID* from EA to EF declined, respectively 0.963, 0.947, 0.931, 0.926 and 0.871.

ID was indeed shown to be a remarkably robust determinant of the mean time within each condition, but the correlation between *MT* and *ID* became much weaker when data from all the operational biases were merged in one regression (see Figure 3). For linear steering, *ID* accounted for only 31.7% of the variance of mean trial completion time caused by both different levels of *ID* and the quite different five operational biases. For circular steering, *ID* could also account for only 44.2%.

3.4.2 Standard Deviation (SD)

ANOVA analysis showed that there was a significant effect of bias ($F_{4, 135}=21.09$, $p<.00001$ for linear tasks, $F_{4, 135}=54.33$, $p<.00001$ for circular tasks) upon standard deviation. Mean standard deviation for EA, A, N, F and EF biases were respectively 1.34, 1.42, 1.64, 1.70, and 2.07 for linear steering and 2.09, 2.41, 2.66, 2.89, and 3.78 for circular steering.

SD varied irregularly with the increase of *ID* for both linear and circular steering tasks. So we further examined the influence on *SD* separately by tunnel amplitude and width.

ANOVA analysis showed a significant effect of amplitude ($F_{2, 135}=19.56$, $p<.00001$ for linear tasks, $F_{2, 135}=11.64$, $p<.00001$ for circular tasks) and width ($F_{2, 135}=32.29$, $p<.00001$ for linear tasks, $F_{2, 135}=138.01$, $p<.00001$ for circular tasks) upon standard deviation. Mean standard deviation for 250, 350, and 450 pixel amplitudes were respectively 1.41, 1.65, and 1.83 for linear steering and 2.55, 2.79, and 2.96 for circular steering. Mean standard deviation for 10, 25, and 40 pixels width were respectively 1.35, 1.66, and 1.89 for linear steering and 1.94, 2.84, and 3.51 for circular steering.

Although there was a significant effect of amplitude on *SD*, it was smaller than the significant effect of width since only 30% and 16% enhancements of *SD* for linear steering and circular steering respectively were observed for the amplitude from 250 to 450 pixels, while 40% and 81% for the width from 10 to 40 pixels. That is to say, only 30 pixel changes in

width resulted in larger enhancement of *SD* (*SD* varied from 1.35 to 1.89 for linear steering, and 1.94 to 3.51 for circular steering), while 200 pixel changes in amplitudes resulted in smaller enhancement of *SD* (*SD* varied from 1.41 to 1.83 for linear steering, and 2.55 to 2.96 for circular steering). So, *SD* was mainly affected by operational biases and tunnel widths (ignoring the smaller effect of amplitude).

According to our analysis, *SD* increased in line with operational biases from EA to EF and with widths from 10 to 40 pixels.

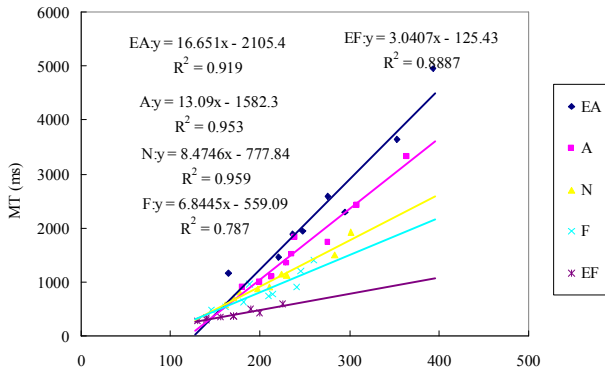
3.4.3 Out of Path Movement (OPM)

ANOVA analysis showed that there was a significant effect of bias ($F_{4, 135}=19.75$, $p<.00001$ for linear tasks, $F_{4, 135}=26.10$, $p<.00001$ for circular tasks) and width ($F_{2, 135}=47.15$, $p<.00001$ for linear tasks, $F_{2, 135}=46.87$, $p<.00001$ for circular tasks) upon *OPM*. Mean *OPM* for EA, A, N, F and EF biases were respectively 0.03%, 0.27%, 0.45%, 1.46%, and 3.73% for linear steering and 0.12%, 0.21%, 0.9%, 1.83%, and 5.7% for circular steering. Mean *OPM* for 10, 25, and 40 pixels were respectively 3.29%, 0.15%, and 0.12% for linear steering and 4.51%, 0.62%, and 0.13% for circular steering.

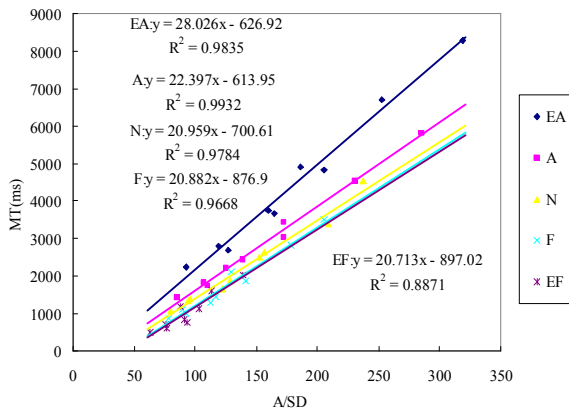
3.5 Model Deduction and Verification

Now, in order to find out a new steering model, we would like to further examine the effects on performance of *MT* with operational biases from EA to EF separately by tunnel amplitudes and widths for both linear and circular steering tasks.

Analysis showed that *MT* decreased in line with operational biases from EA to EF and with width from 10 to 40 pixels for both steering tasks, which had opposite changes tendency compared with *SD*, while increased in line with amplitude from 250 to 450 pixels for both steering tasks.



(a) Linear steering task



(b) Circular steering task

Fig. 4 Mean MT vs. A/SD for each bias in both linear and circular steering tasks

According to foregoing analysis, it is presumably predicted that more *MT* is needed with the increased tunnel amplitude, while less *MT* is needed with the increased *SD*. So, the following hypothetic equation is given:

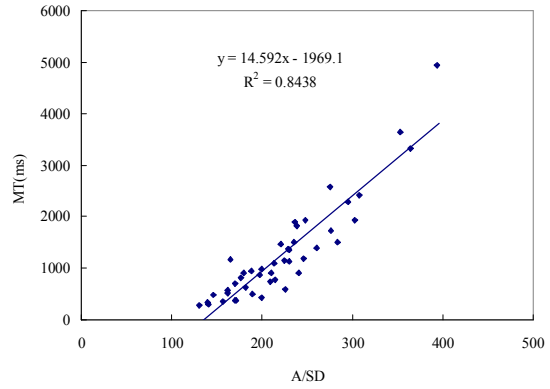
$$MT = a + b * ID_n$$

Where, ID_n is a new index of difficulty and formulated as:

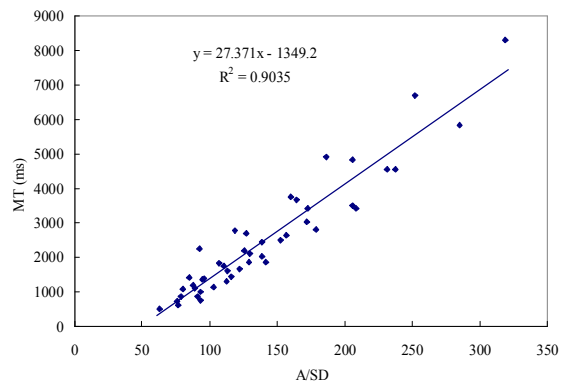
$$ID_n = A/SD$$

SD is the standard deviation of sampled points.

Next, we would examine the predictive power of this newly proposed model for both linear and circular steering tasks in each operational bias and across all the operational biases (see Figure 4 and Figure 5 respectively).



(a) Linear steering



(b) Circular steering

Fig. 5 Mean MT vs. A/SD for all biases combined in both linear and circular steering tasks

From Figure 5, we could clearly observe that the new index of difficulty ID_n was a stronger determinant than ID when data from all conditions were merged in one regression. ID_n could account for 84.4% and 90.4%, respectively for linear and circular steering, of the variance of mean trial completion time caused by both different levels of ID_n and the quite different five operational biases.

An interesting discovery is that the R^2 values of *MT* vs. ID_n linear regression in biases A, N and EF were even higher than their corresponding R^2 values of *MT* vs. ID in the same bias in linear steering tasks. In circular steering tasks, the R^2 values of *MT* vs. ID_n linear regression in biases EA and A were higher than their corresponding R^2 values of *MT* vs. ID in the same bias (see Figure 4).

3.6 Conclusions

We have systematically explored the effect of different operational biases of subjects (toward to speed or accuracy) on steering tasks for both straight and circular movements. Experimental results showed that the effect of subjective factors indeed existed. Different operational biases would result in different levels of *SD*, which was mainly affected by

the different operational biases of subjects and by tunnel widths. Then, we deduced a new steering model involving system and subjective factors.

Two interesting discoveries in our investigation were, firstly, the effect of subjective factor indeed existed in steering tasks, which was reflected by different levels of *SD* and *MT*; secondly, our newly proposed model was still shown to be a robust determinant of the mean steering time within each operational bias for both linear and circular steering tasks. When all the operational conditions were merged in one regression, the new model was shown to be a much more predictive determinant of the mean steering time than the traditional steering model.

References

- (1) Card, S. K., Moran, T. P., and Newell, A. The Keystroke-level Model for User Performance Time with Interactive systems. *Communications of the ACM*, 23, (1980) 396-410
- (2) Fitts, P. M. The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *J. of Experimental Psychology*, 47, (1954) 381-391
- (3) Guiard, Y. Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as A Model. *Journal of Motor Behavior*, 19, (1987) 486-517
- (4) Epps, B. W. Comparison of Six Cursor Control Devices Based on Fitts' Law Models. *Proceedings of the Human Factors Societies 30th Annual Meeting, CA: Human Factors Society*. (1986) 327-331
- (5) MacKenzie, S., Sellen, A., and Buxton, W. A Comparison of Input Devices in Elemental Pointing and Dragging Tasks. *Proc. CHI'91*, (1991) 161-166
- (6) MacKenzie, I., and Zhang, S. The Design and Evaluation of A High-performance Soft Keyboard. *Proc. CHI'99*, (2005) 25-31
- (7) Accot, J., and Zhai, S. Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks. *Proc. CHI'97*, (1997) 295-302
- (8) Accot, J., and Zhai, S. Performance Evaluation of Input Devices in Trajectory-based Tasks: An Application of the Steering Law. *Proc. CHI'99*, (1999) 466-472
- (9) Accot, J., and Zhai, S. Scale Effects in Steering Law Tasks. *Proc. CHI2001*, (2001) 1-8
- (10) Dennerlein, J., Martin, D., and Hasser, C. Force-feedback improves performance for steering and combined steering-targeting tasks. *Proc. CHI2000*, (2000) 423-429
- (11) Pastel, R. Measuring the Difficulty of Steering Through Corners. *Proc. CHI'06*, (2006) 1087-1096
- (12) Zhai, S., Accot, J. and Woltjer, R. Human Action Laws in Electronic Virtual Worlds - an empirical study of path steering performance in VR. *Presence*, 13(2), (2004) 113-127
- (13) MacKenzie, I. Motor Behavior Models for Human-Computer Interaction. J.M.Carroll (Ed.) *HCI model, theories, and frameworks: Towards a multidisciplinary science*, (2003) 27-54
- (14) Zhai, S., Kong, J., and Ren, X. Speed-accuracy tradeoff in Fitts' law----on the equivalency of actual and nominal pointing precision. *Int. J. Human-Computer Studies*, 61(6), (2004) 823-856
- (15) MacKenzie, I. Fitts' Law as a Research and Design Tool in Human-computer Interaction. *Human Computer Interaction*, 7, (1992) 91-139
- (16) Ren, X., Kong, J., and Jiang, X. SH-Model: A Model Based on Both System and Human Effects for Pointing Task Evaluation. *IPSI Journal*, 46(5), (2005) 1343-1353
- (17) MacKenzie, I. A Note on the Information-theoretic Basis for Fitts' Law. *Journal of Motor Behavior*, 21, (1989) 323-330

- (18) Kulikov, S., MacKenzie, I., and Stuerzlinger, W. Measuring the Effective Parameters of Steering Motions. Proc. CHI2005, (2005) 1569–1572.
- (19) Fitts, P. M., and Radford, B.K. Information Capacity of Discrete Motor Responses under Different Cognitive Sets. Journal of Experimental Psychology, 71, (1966) 475–482

4. Research achievement in image recognition algorithm area-1: Motion

We summarize the research achievement in image recognition algorithm area-1: Motion by one article that was published by Tomofumi Tanaka, Yuji Nagano and Mamoru Okada as the proceedings of NEINE '07 (ref.14 in 1.3). The paper is shown below.

Yubimoji recognition by distance evaluation.

Tomofumi TANAKA, Yuji NAGANO and Mamoru OKADA

Abstract

This research aims to recognize yubimoji by the distance evaluation method. Yubimoji is a system of manual kana used as part of Japanese Sign Language. It is a signary of 45 signs and 4 diacritics representing the phonetic syllables of the Japanese language. Manual size and the speed of expressing a word by yubimoji are different by a person. Therefore the recognition method that is hard to influence them is needed. The movement of expressing yubimoji is followed by using two-dimension measurement software, and this research is aimed at inspecting about recognition method of yubimoji by its data.

4.1 Introduction

Hearing impairment is a full or partial decrease in the ability to detect or understand sounds. Hearing impaired person communicates by Sign Language that is a language which uses manual communication, body language and lip patterns instead of sound to convey meaning. In Japan, Sign Language interpreter are much fewer than Sign Language native speaker. Therefore, it is difficult for many people to communicate with a

hearing impaired person smoothly. If computer recognizes Sign Language, we would be able to communicate with hearing impaired person more smoothly. Yubimoji is a system of manual kana used as part of Japanese Sign Language. It is a signary of 45 signs and 4 diacritics representing the phonetic syllables of the Japanese language. If Sign Language speaker don't know word and proper noun how to express by Sign Language, yubimoji can express them. This research aims to recognize yubimoji.

4.2 Taking yubimoji picture

The subjects for taking yubimoji picture are five people, and stereo-camera was used. Recognition objects were assumed "a", "i", "u", "e", "o", "te" and "ne", because "a", "i", "u", "e" and "o" are a vowels (FIGURE 1) and forms of yubimoji of "te" and "ne" resemble each other. Yubimoi picture was taken from the front for taking all finger-tips. The subjects refer to a book and performed yubimoji.

4.3 Motion capture

Each finger-tips and wrist like FIGURE 2 is followed by motion capture at the movement of yubimoji (FIGURE3), and the distances from each finger-tip B - F to the wrist point A are calculated.

4.4 The distance evaluation method of 3 evaluations

First, in the distance evaluation method of 3 evaluations, the distances from each finger-tip B - F to the wrist point A are calculated. The distance that is the farthest from A is given evaluation 3 (FIGURE 4), the distance that is the nearest to A is given evaluation 1 (FIGURE 5) and the intermediate distance is given evaluation 2.

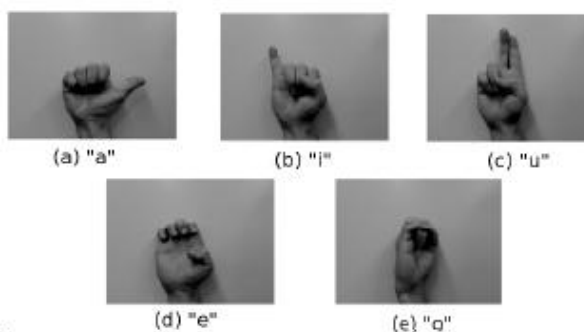


Figure1: Example of Yubimoji

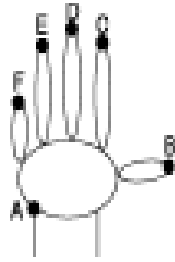


Figure2: Followed Point



Figure 3: Motion Capture

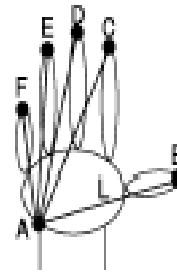


Figure4: Distance Evaluation 3



Figure5: Distance Evaluation 1

The intermediate distance is calculated by expression (1). In the expression (1), M denotes intermediate distance, L denotes the farthest from A and l denotes the nearest from A.

$$M = \frac{L + l}{2} \quad (1)$$

Their data are compared with the distance of each finger-tip of yubimoji to want to recognize, and the differences of distance of each finger-tip are calculated. Next, all the fingers are given distance evaluation 1, 2 or 3 from difference of that distance. This method recognizes yubimoji by using combination of that evaluation.

4.5 Advantages of the distance evaluation method of 3 evaluations

Advantages of this method are as follows :

1. The method that needs sample of all yubimoji takes many times. But this method needs only data of FIGURE 4 and 5 of human that perform yubimoji for recognizing yubimoji. Therefore this method can omit trouble to take sample of all yubimoji.
2. This method don't depend on the speed of performing yubimoji.
3. This method don't depend on manual size.

Table 1. Evaluation of distance

| | |
|--------------|---------------------|
| Evaluation 3 | the farthest from A |
| Evaluation 1 | the nearest to A |
| Evaluation 2 | the intermediate |

4.6 Result and consideration

The "a", "i", "u", "e" and "o" were recognized by the distance evaluation method of 3 evaluations. But "te" and "ne" were not recognized by this method because their combinations of evaluation were same (FIGURE 6). So their coordinate data was used because the position of wrist point and the middle finger-tip turns over. And "te" and "ne" were recognized by comparing the coordinate data of middle finger-tip with the coordinate data of wrist point. As a result, "a", "i", "u", "e", "o", "te" and "ne" were recognized.

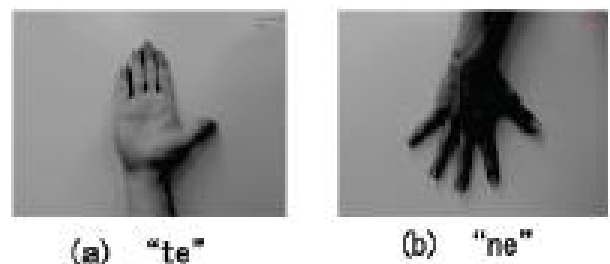


Figure6: "te" and "ne"

4.7 Conclusion

In this research, the yubimoji recognition has been performed by the distance evaluation of 3 evaluations. And this method was able to recognize seven kinds of yubimoji. But there is a possibility that the intermediate distance that is average of the farthest distance and the nearest distance would not be evaluated by degree of bending of finger. Therefore, it is necessary to examine a method that gives redundancy to the intermediate distance. A two-dimensional measurement may misrecognize because distance changes depending on camerawork. Therefore a three-dimensional measurement is necessary. As a future problem, it is necessary to examine whether the distance evaluation method of 3 evaluations can recognize all yubimoji.

Reference

1. Kiyoshi Tanaka, "Let's speak by sign language", Seitousya, 2001.(In Japanese)
2. Kazuhito Fujimura, "Method of Extracting Feature Frame by 3-Dimensional Measurement from Sign Language Video", KUT, 2006.(In Japanese)

5. Research achievement in image recognition algorithm area-2:Image retrieval

We summarize the research achievement in image recognition algorithm area-2: Image retrieval by one article by Yoshimasa Kimura as the summary of research. The paper is shown below.

Similar Image Retrieval Using Multidimensional Psychological Space. *Yoshimasa KIMURA*

Abstract

Information treated by the computer increases rapidly when surrounding network develops. Especially, when a huge amount of images gather, the presentation technology that makes easily to see by arranging the image is required. Moreover, the demand to retrieve a specific image is caused. We examine the technology which retrieves similar images to the presented query image as similar image retrieval. We propose the image navigation which repeats narrowing the image over the

multistage on the multidimensional psychology space which reflects subjectivity evaluation value of the non-similar degree between images. This paper shows the effect of this technology and show whether the axis of the multidimensional psychology space refers any feature on the multidimensional psychology.

5.1 Introduction

When the network develops, a lot of images are taken into the computer. The image retrieval becomes an important problem. The usual image retrieval is executed by using the key word. This is efficient when there are a lot of different kinds of images. However, when a target image is similar image, it is difficult to retrieve by using the key word. There is similar image retrieval in the field of the image retrieval. At this time, the retrieval system finds desired image by using given image.

When human sees the image, he obtains various senses and impressions. However, these have information with high dimension, and the element of the dimension is different depending on the person. If the retrieval system can extract the essence from higher dimension, the system can efficiently retrieve the image that the user desires. However, even if the essence can be extract from high dimensional information, the human can't see the essence.

To solve this problem, there is a concept of the information visualization. As the method to realize it, there are a self-organizing maps[1], a multidimensional scaling[2] and a principal component analysis, etc. The self-organizing maps and the principal component analysis need the feature of the image. However, the feature does not necessarily agree to human's sense.

This paper describes similar image retrieval by image navigation using the multidimensional scaling to match retrieval result to human's sense. The pictorial navigation[3] originally offers management mechanism of the image information that the user can correct the intention of the retrieval little by little while thinking. In this paper, the pictorial navigation is applied to the similar image retrieval. The proposed method retrieves the image over the multi-stage by changing the axis obtained by multidimensional scaling. The operation that repeats narrowing the image is a new point.

5.2. Image navigation

This chapter describes the method of applying the pictorial navigation to the similar image retrieval and explains the construction of retrieval system and use method of the system.

5.2.1 Multidimensional scaling

Multidimensional scaling arranges the relations between individuals in the spaces of two or three dimensions. The method has a merit that can decide arrangement of data in multidimensional space while maintaining the size relation of the similar level between data.

5.2.2 Principle of similar image retrieval by multidimensional scaling and image navigation

Figure 1(a) shows the multidimensional psychology space, and the image that human saw exists in the space. Figure 1(b) shows two dimensional screen on the computer, and the image is arranged on this. The method that transform from multidimensional space to retrieval screen is needed. There is the multidimensional scaling as a method for enabling it. The subjectively evaluation value that shows the non-similarity between each image is necessary for using multidimensional scaling. The retrieval system acquires the axis with k dimension by applying multidimensional scaling to the subjectively evaluation value.

The system arranges the image on the retrieval screen by using the two axes. The system can retrieve images near the user's specification and narrows the number of images. The system can decrease the number of retrieved images by repetition of arrangement the retrieved images and its narrowing.

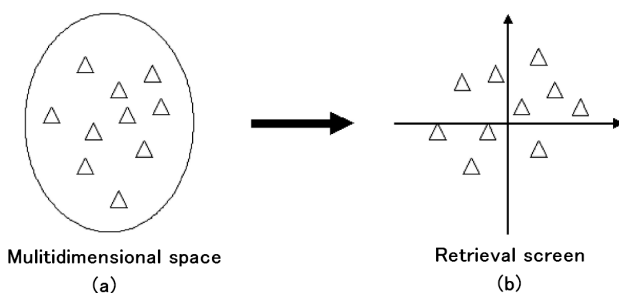


Fig. 1 Transform from multidimensional space to two dimensional space.

5.2.3 Construction of similar image retrieval system

The image navigation performs an efficient retrieval of a similar image. The system has two functions. The first is to accept subjectively evaluation value and output the result obtained by the processing of multidimensional scaling. The second is to select the image located internally within radius r . Where, coordinates at centered position or centered image and retrieval radius r are given by the user. The construction procedure of the system that realizes these two functions is as follows.

- Step1 Make the data that the subjectively evaluation value between images was stored. Where, the value is obtained by inputting the value on the screen as shown in Fig. 2.
- Step2 Apply the multidimensional scaling to the data, and acquire k axes and its coordinate value for each image.
- Step3 Make the program that selects two axes whose variance of score obtained by multidimensional scaling is maximum and second maximum among k axes and arranges each image on the two dimensional plane composed of the two axes.
- Step4 Make the program that select the image located internally within radius r which is given by the user.
- Step5 Compose the program that can repeat from Step2 to Step4.

5.2.4 Use method of the system

When the query image is given to the user, the user sees the query image and retrieves the image by using

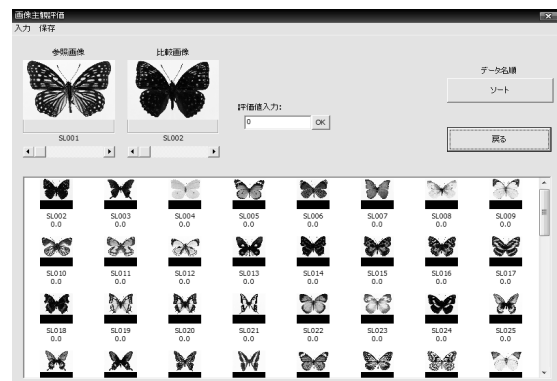


Fig. 2 Input screen of subjectively evaluation value.

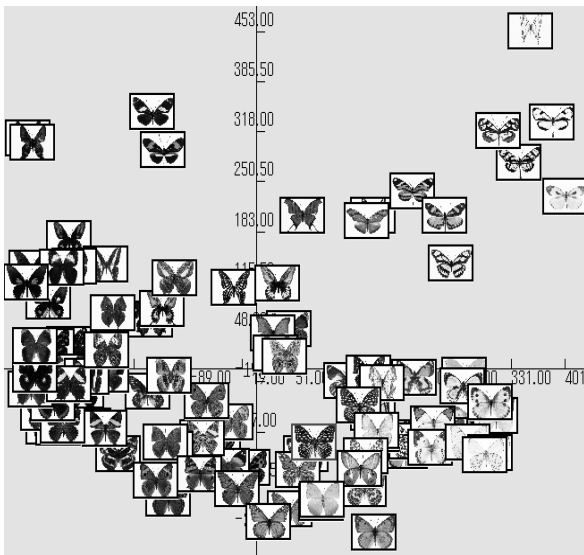


Fig. 3 Retrieval screen

this system. The results are shown as follows. Figure 3 shows the retrieval screen.

- Step1 The image navigation system selects two axes from among k axes obtained by multidimensional scaling. Images to be retrieved are displayed on the two dimensional space defined by the two axes.
- Step2 The user finds the image which is similar to the query image and inputs the coordinates at center to be retrieved and retrieval radius r , then, obtains the retrieval result.
- Step3 When the user intends to narrow images further, the user executes Step1 and Step2 for the retrieval result. When the user judged that the image near query image is retrieved, the system completes processing.

5.3. Experiments

To test the efficiency of the proposal method, experiments using butterfly specimen image were made.

5.3.1 Image data

The butterfly images on the marketed were used for an experimental image. The number of sheets N of the image is $N = 100$. Figure 4 shows the examples of butterfly's image. Butterfly's color includes black, gray and brown, etc., and image includes butterfly's whole body.

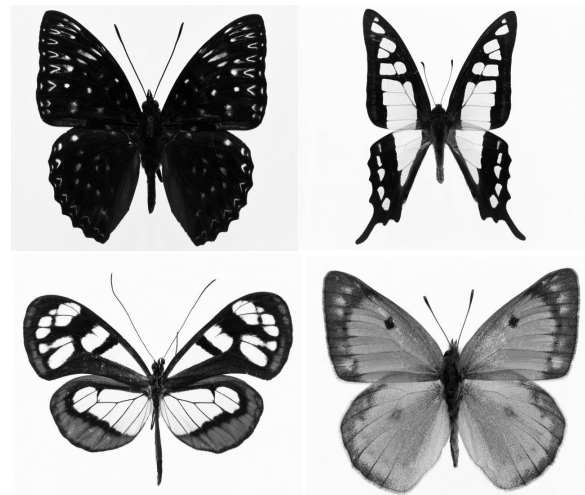


Fig. 4 Examples of butterfly's image

5.3.2 Making of the image navigation system

The image navigation system was constructed by the method described in 2.3 by using $N = 100$ and $k = 20$. It was confirmed that the system performed well as expected.

5.3.3 Experimental method

Four query images and 20 axes obtained by the multidimensional scaling were used. Subject retrieves the image which is similar to the query image. The experiments were conducted by the following three groups. Subjects are four peoples in each group. The subjects see all images, and select the image near query image by using the system. Group A, B, and C retrieve the image by using following two axes.

- Group A Two axes chosen in significant order from among 20 axes by the system.
- Group B Two axes chosen by random numbers by the system.
- Group C Two axes selected by subject's intuition.

5.3.4 Evaluation method

The experimental result is evaluated from the following three viewpoints.

(1) Retrieval inclusion rate

Retrieval inclusion rate α is defined by $m / n \times 100$, where m and n are defined as the number of retrieved

similar image and the number of the retrieved images, respectively. The larger the value of α is, the better the performance of the system is.

(2) Undetection rate

Undetection rate β is defined as a ratio of the number of the similar images that remain without being detected to the number of similar image. The smaller the value of β is, the better the performance of the system is.

(3) Frequency of narrowing image

Frequency of narrowing image γ is defined as the number of screen until desired image is found. The smaller the value of γ is, the better the performance of the system is.

5.3.5 Experimental results

Table 1 shows comparison of retrieval performances in group A, B, and C. As for (1), retrieval inclusion rate α of the Group A, B and C were 60% , 46%, and 63%, respectively. For (2), undetection rate β of the Group A, B and C were 56%, 68%, and 63%, respectively. For (3), frequency of narrowing image γ were 2.3, 2.7, and 2.7, respectively. Evaluation values of group A and group C are better than that of group B in all of the three items.

Table. 1 Comparison of retrieval performances in group A, B, and C.

| Item \ Group | (1) α (%) | (2) β (%) | (3) (times) |
|--------------|---------------------|--------------------|----------------|
| A | 60 | 56 | 2.3 |
| B | 46 | 68 | 2.7 |
| C | 63 | 63 | 2.7 |

Comparing group A and group C, the performance of α is almost same. The result shows there is little difference when axis with significant is used and when the axis that the user had selected is used. It was proven the retrieval used multidimensional scaling is suitable for human's sense. Moreover, the result draws the method using axis with significant is effective because the labor of the operation is fewer for the method than the case that the human chooses the axis.

5.4. Discussion

The experimental result in 3.5 showed the effectiveness of the proposed system. However, the meaning of two axes led by multidimensional scaling is not clarified. Two axes seem to have the psychological factor. On the other hand, the quantity of a physical feature can be extracted from the image. If the relation between the axis of multidimensional scaling and the quantity of a physical feature is obtained, the fact is useful for the clarification of the meaning of the axis.

The color and shape are used as physical feature. As the color feature and shape feature, the color histogram with 10 dimensions[1] and the peripheral feature[2] with 20 dimensions were adopted respectively. Assigning the axis of multidimensional scaling as objective variable and physical feature as explanatory variable, the relation between both is obtained by a multiple regression analysis.

A multiple regression analysis was made using 50 images of the butterfly. Coefficients of determination R^2 of axis 1 and axis 2 are 0.821, 0.854, respectively, and the result of the analysis can be reliable enough. The feature corresponding to the explanatory variable that the partial regression coefficient becomes the maximum in axis 1 and axis 2 is a proportion of black and width of the rear wing, respectively. It is observed that the more it goes to the left the more the number of black butterfly increases in axis 1 of Fig. 3, and the more it goes to the lower side the more the number of butterfly with long width of the front wing increases in axis 2 of Fig. 3. This result agrees with the result of viewing in Fig. 3 well. As a result, which characteristic of the butterfly the axis of multidimensional scaling referred became clear.

5.5. Conclusion

This paper proposes the image navigation which uses the multidimensional psychological space to perform the similar image retrieval suitable for human's sense. Human's sense to image is given as subjectively evaluation value. Applying the multidimensional scaling to the image, the retrieval system acquires k axes and its coordinate value for each image. The images are displayed on two dimensional screens by using two axes among the k axis. The user retrieves desired image to perform image navigation that the user

repeats retrieval and to narrowing the retrieval image. The performance and efficiency of image retrieval can be expected to improve by these measures.

On the experiments using the butterfly's similar image, it was proven that the retrieval using multidimensional scaling is suitable for human's sense and the lover of the operation is few because it selects the axis automatically. The experimental results draw the proposed method using multidimensional scaling is efficient.

References

- [1] T. Hyuga, I. Nishikawa, "Implementing the database system of butterfly specimen image by Self-Organizing Maps," *Journal of Japan Society for Fuzzy Theory and Systems*, vol.14, no.1, pp.74-81, Feb. 2002.
- [2] K. Tanabe, J. Ohya and K. Ishii, "Similarity retrieval method using multidimensional psychological space," *IEICE Trans. Inf.& Syst.*, vol.J75-D-□, no.11, pp.1856-1865, Nov. 1992.
- [3] H. Koike, H. Kasahara, "Implementing the database system of butterfly specimen image by self-organizing maps," *Proceeding of the IEICE*

General Conference, D-389, 1990.

6. Future work

In order to enhance research progress in this project, we have introduced some Ph.D. students in Special Scholarship Program in KUT and some of master students. With these graduate students, research themes shown above will be continued and also, we are expecting to make new research themes in applicable human perception area. We hope achievements will be more and better and also, we are expecting to establish some researches in image recognition algorithm area, starting in 2008 academic year soon.

Acknowledgement

This work was supported by the Special Subsidies in Subsidies for ordinary expenses of private schools from the Promotion and Mutual Aid Corporation for Private Schools of Japan.

サラウンディング・コンピューティングにおける人間中心視覚 情報利用研究プロジェクト 2007年度報告

篠森敬三* 岡田 守** 木村義政† 任 向実***

高知工科大学 工学部 情報システム工学科

〒782-8502 高知県香美市土佐山田町宮ノ口 185

E-mail: * shinomori.keizo@kochi-tech.ac.jp, ** okada.mamoru@kochi-tech.ac.jp,
*** ren.xiangshi@kochi-tech.ac.jp

要約: 視覚情報は人間に対する情報提示手法として最も重要であり、サラウンディング・コンピューティング (SC) を実現するためにも包括的な知識や応用可能なアルゴリズムが必要である。

人間の知覚・作業速度については、まず、極短時間光への視覚系全体の時間的応答であるインパルス応答関数 (IRF) による S 錐体経路の時間的特徴について調べた。結果は、S 錐体オン IRF は輝度の IRF より遅かったが、オフ IRF より速かった。さらに、S 錐体オンとオフの IRF の違いは、統計学的に有意であった。さらに、ペン操作での軌道運動タスク遂行時間における主観的バイアス効果について調べた。従来は運動時間とシステム要因 (パス幅と振幅) のみを定式化していた。遂行時間と標準偏差に影響を与える主観的な操作バイアス (操作速度と正確さ) の効果を調べる実験に基づいて著者が提案したモデルは、より正確な予想を可能とした。

動画と画像の自動処理については、指と手首の距離評価手法を利用した手話指文字の自動認識処理を行った。指文字は、手で仮名を作るものであり、45 のサインと 4 つの発音区別符号からなる。指文字では、手のサイズと表示速度が手話者によって異なるため、それらに影響されない認識方式が必要であった。そこで指文字を表す運動を二次元距離測定により追従し、距離データ変化から指文字認識を行った。また、人間の感覚にあった類似画像を検索するため、画像間の非類似度の主観評価値を反映した多次元心理空間上で画像の絞り込みを繰り返す手法を提案し、その効果と多次元心理空間軸での特徴を示した。

†現在の所属機関は崇城大学情報学部コンピューターシステムテクノロジー学科 (E-mail: kimura@cis.sojo-u.ac.jp)。