1. INTRODUCTION

Interactive evolutionary computation (IEC) is used to support product development. This method uses user Kansei (subjective) evaluation information for finding candidate solutions of evolutionary computation (EC) instead of the evaluation function of the EC method. The IEC system is effective for creating objects using the Kansei information of users. Although users require functional products, IEC systems help designers in developing products that can also satisfy the subjective and emotional preferences of users [1]. IEC systems include image retrieval [2], music generation [3], color coordinates [4], room layout [5], 3D game model generation [6], etc.

A traditional IEC-based system suffers from large user evaluation loads of candidate solutions. To solve this problem, researchers have proposed to simplify the user evaluation methods or to apply the biological information of users in their evaluation of candidate solutions. Using a simple evaluation method, previous studies have proposed a tournament evaluation method that uses a paired comparison method [7] or an interactive tabu search method where users select their favorite candidate solutions from an array of proposed solutions [8]. Although these methods are effective in reducing user evaluation loads [7, 8], they require users to evaluate candidate solutions expressly, which entails them to cope with a certain evaluation load.

In contrast, by exploiting the user’s biological information, such as heartbeats or brainwaves, the system can obtain user evaluation information while the user is viewing the candidate solutions. Previous studies have verified the effectiveness of this method for obtaining candidate solutions that satisfy user preferences [9, 10], and have achieved user evaluation without evaluating the candidate solutions expressly. However, users must wear special devices, such as heart rate meters or electroencephalographs, for their biological information to be measured, which has contributed to the difficulty in popularizing these IEC systems.

To solve these problems, we apply user gaze information in the evaluation of candidate solutions in IEC. Gaze information includes the user’s potential preferences, which are derived from various processes. When user gaze information is applied in the evaluation of candidate solutions, IEC systems can obtain user evaluation information while users are viewing multiple candidate solutions. In this paper, we verify the effectiveness of the eye tracking IEC system using evaluation experiments with real users. In the experiment, we use a normal IEC system as a comparison method where users manually evaluate candidate solutions using a 10-stage evaluation process. The experimental results show that the eye tracking IEC method can generate solutions with results equivalent to those of the compared system.

Keywords: Gaze information, Interactive evolutionary computation, Eye tracker
a color coordination system with one-max optimization using an eye tracker [12]. Holmes and Zanker also proposed an IEC system with eye tracking that optimizes the rectangular aspect ratio [13]. Although the results of these studies achieved a certain level of effectiveness, they failed to clearly verify the overall effectiveness of a user gaze information-based method. Moreover, user motivation is considered an issue in the evaluation of objects used in these systems.

On the other hand, García-Saravia et al. attempted to use gaze information for the evaluation method of the IEC system [14]. However, they only analyzed the correlation between user gaze duration information and the evaluation made by their system parameters and did not employ user gaze information for evaluating candidate solutions of IEC. Moreover, Jung et al. proposed to employ user gaze information for the recommendation of user preference items for online shopping [15]. However, they did not evaluate user evaluation loads. In this study, we investigate user evaluation loads while using the IEC system with gaze information.

Previously, we applied user gaze information for evaluating candidate solutions of IEC and verifying the effectiveness of the proposed system by user experiments [16]. In this study, we investigate more effectiveness of the method with gaze information in terms of the evolutionary performance of candidate solutions. In this system, we use a genetic algorithm (GA) within an evolutionary algorithm of IEC.

We conduct an evaluation experiment to verify the effectiveness of the proposed method, namely eye tracking IEC (E-IEC) and normal IEC (NIEC). In the first method, users can view the presented designs by using the system. In the second method, they manually evaluate all designs using a 10-stage evaluation process. The experiment uses the NIEC method for comparison where users evaluate candidate solutions using a 10-stage evaluation process and demonstrate the satisfaction levels of the system-generated designs and user evaluation loads.

2. E-IEC SYSTEM

2.1 Schematic of E-IEC system

Figure 1 shows the schematic of the E-IEC system. First, the system generates the initial gene types and consists of phenotypes (candidate solutions that are then presented to the user) corresponding to each gene types. Then, the user views all candidate solutions, while the system measures the user gaze information. When the user finishes evaluating the candidate solutions, the system obtains user evaluations. Finally, the system operates GA processing, which results in the generation of new gene types and phenotypes that are again presented to the user. The system repeats these operations and creates a user preference design. In the E-IEC system, each candidate solution is evaluated using user gaze information. We use a Tobii Pro X2-30 eye tracker to obtain user gaze information. This device can measure user gaze information at a rate of 30 [Hz].

Figure 2 shows the candidate evaluation of E-IEC. The E-IEC system divides the interface into eight areas. The E-IEC system first measures user gaze positions for each sample (black circles in Figure 2) and then counts the number of user gaze positions for each candidate solution. If the gaze position is out of the interface, the system does not measure and count the number of user gaze positions.

For example, when the number of user gaze position samples of each candidate solution is $s_1 \sim s_8$, each $s_i$ is equal to 3, 2, 2, 1, 3, 2, and 2. Next, the E-IEC system normalizes these $s_i$ in a range of 0-100. Then, the evaluation values of each candidate solution, $v$, are 100, 50, 50, 0, 100, 100, 50, and 50. The system uses these evaluation values $v$ for GA processing.

The proposed system decides the evaluation values of all candidate solutions based on the number of gaze samplings. However, users may view the design only because their interest. Then, a design that has more gaze samplings may not be considered as user preference. In this case, the proposed system may obtain inaccurate evaluation information (noise) on user preferences and generate designs.
that do not satisfy users. Therefore, we demonstrate that
the proposed system can generate designs based on user
preferences even if the user evaluations include inaccurate
information from prior experiments.

2.2 Candidate solutions to our system

We use running shoes designs as IEC evaluation objects
similar to our previous system [16].

Figure 3 shows the gene coding of the running shoes
design system, which consists of the following five parts:
sole, base, toe, line, and shoelace. The system creates various
running shoes designs by recombining parts of each design.

Figure 4 shows the parts of running shoes design. Eight
designs are available for each part and are expressed by
three bits. Therefore, the total gene length is 15 bits and
the system can generate 32,768 running shoes designs.
We assign a bit pattern to each design considering the
similarity between the design appearance of each part and
gene rows.

![Figure 3: Gene coding of the running-shoes design](image)

![Figure 4: Design parts of the running shoes design](image)

2.3 User evaluation

Figure 5 shows the interface of the E-IEC system.
First, the user starts the system and evaluates the
presented designs by viewing them. The system receives
user gaze information while the designs are being viewed.
When the user has finished viewing the designs, he/she
then clicks the “Next” button and new designs are
presented. The user repeats these operations until his/her
preferred designs are generated. Finally, when the
user has finished using the system, he/she clicks the
“End” button.

Figure 6 shows the experimental setting of the E-IEC
system. The user sits in front of the screen and views
the presented designs. Conversely, in the NIEC system
(the compared system), the subjects input their
evaluation values of the candidate solutions using a
mouse.

3. EVALUATION EXPERIMENTS

3.1 Outline of the experiment

We perform an evaluation experiment to verify the
effectiveness of the E-IEC method from the view point
of evolutionary performance of candidate solutions and
the evaluation loads of users. We employ 16 university
students in their 20s (6 males and 10 females) as subjects. In the experiment, the compared system also uses normalized evaluation values of the 10-stage evaluations for GA operations.

Moreover, the subjects use the E-IEC system in two patterns (E-IEC 1 and E-IEC 2). In this experiment, we assume various use situations of the E-IEC system. For the E-IEC 1, the subjects see as many of their favorite designs as they wish after viewing all designs freely. In the situation, a user expressly evaluates designs using his/her gaze. For the E-IEC 2, the subjects see all designs freely during their evaluation. In the situation, a user only freely evaluates designs using his/her gaze (nearly real use situation). The subjects are considered to have finished using all systems when 80% of the presented designs are viewed as convergence.

Table 1 shows the experimental parameters. The subjects use each system based on the concept that customers will select running shoes designs that they want to wear to enjoy running. We set the mutation rate parameters to 20% to compare high mutation rate with EC method. A high mutation rate occurs in random searching by breaking the high-quality gene information. However, using a random factor (a high mutation rate) in the IEC system, the system can generate or search unexpected suitable designs for the user.

After the experiment, we use questionnaires to survey the satisfaction levels of the generated candidate solutions and usability (i.e., the level of evaluation load) for each system. We measure the time and number of generations by users evaluating the designs for each system. Moreover, we consider the effect of the order in which each system is used. Therefore, the order of use of the three systems is set randomly for each subject.

### 3.2 NIEC system

We use the NIEC system, which has similar interface, as a comparison method in the experiment. However, when inputting the evaluation values of each design, the subjects use the slider bar under each design with a mouse. Then, they input evaluation values for each design. When a user has finished using the system, they click the “End” button.

The NIEC system obtains user preferences by using a 10-stage evaluation process and normalizes these evaluation values in a range of 0 to 100. These normalized values are used for GA operations.

### 4. EXPERIMENTAL RESULTS

#### 4.1 Satisfaction level of generated designs

Figure 7 shows the average and standard deviations of the satisfaction levels of the generated designs from each system. The satisfaction levels of the two systems except E-IEC 2 are approximately 4, whereas E-IEC 1 can generate designs that satisfy the users. We perform the Friedman test to confirm the statistically significant differences in the satisfaction levels of each system. The Friedman test showed that when the significance level was 5%, the statistically significant differences in the satisfaction levels of each system could not be confirmed. Therefore, we confirmed that the three systems could generate designs that satisfied the preference of the subjects. Moreover, the satisfaction level of E-IEC 1 is higher than that of E-IEC 2 because in E-IEC 1, the subjects see as many of their favorite designs as they wish. Then, E-IEC 1 obtains more accurate user evaluation information and well-evolved candidate solutions.

We did not confirm the statistically significant difference among the three method. However, the satisfaction level of two E-IEC systems are smaller than that of NIEC system. This is because that user views the presented designs freely or by simply interested. In addition, designs of the center of evaluation interface are easy to acquire samplings of use gaze information (high evaluation value) because user eye movements often cross designs in the middle of the screen when user views designs that placed at both ends of the interface. Then, the eye tracker ends up counting user gaze samplings even if eye movement of the left and right, and includes some noise data.

To avoid miscount, the system adds each evaluation value of design when gaze samplings are measured n times a row. In this way, we consider that a longer staying of the use gaze shows that the user likes the viewed design. However, this method cannot remove influence that the user views designs by simply interested.

Figure 8 shows the distribution of satisfaction levels. Approximately 60% of all subjects assign a value of 4 to the generated designs independent of the system. Moreover, the satisfaction levels of E-IEC 2 are smaller than those of E-IEC 1 and NIEC because the subjects can view the designs freely during their evaluation. Moreover, the system has difficulty in measuring the superiority of a particular design.
In addition, we discuss the designs created by the subjects to survey user preference and satisfactory designs. Figure 9 shows the results of the created designs that are assigned a satisfaction level of four or five. The single star (*) and double stars (**) in the upper left of some designs represents designs generated using the E-IEC 1 and 2 systems. We confirm that 60% of these designs are generated using E-IEC systems. In addition, 55% (45%) out of those were created by the E-IEC 1 and 2 systems. Therefore, we confirmed that the designs that satisfy users include the created designs by E-IEC systems.

The results show that male subjects create similar simple designs, while female subjects generate various colorful or simple designs. Therefore, these results show the difference between males and females in the creation of running shoes designs.

4.2 Usability of the systems

Figure 10 shows the average and standard deviations of the usability of each system. The usabilities of E-IEC 1 and E-IEC 2 were higher than that of NIEC. While using NIEC, the subjects had to evaluate the designs manually. However, in the E-IEC system, the user can easily evaluate the designs as they view them. The operation loads of the E-IEC system are smaller than those of NIEC. Moreover, the usability of E-IEC 2 is slightly higher than that of E-IEC 1 because subjects using the E-IEC 2 system are viewing the designs freely. Finally, the psychological loads of the subjects when using E-IEC 2 are smaller than those when using E-IEC 1.

We perform the Friedman test to confirm the statistically significant differences in the usability of each system. The Friedman test shows that when the significance level is 5%, the statistically significant differences in the usability of each system are confirmed. Moreover, Scheffe’s multiple comparison results show that when the significance level is 5%, the usabilities of E-IEC 1 and 2 are significantly higher than that of NIEC.

Figure 11 shows the usability distribution. Approximately 80% of all subjects assign a value of 4 or 5 to the
usability for E-IEC 1 and 2. Then, E-IEC 1 and 2 are effective in reducing the evaluation loads of users. However, some subjects commented that “he/she feels the fatigue with the E-IEC systems because they see the presented designs with a fixed stare.” The fixed stare has no difference in the normal gaze for measuring the gaze information. Thus, we will revise the experimental conditions. As a result, the E-IEC system is effective in reducing the evaluation loads of users and can generate designs suitable to the user’s preferences.

### 4.3 Evaluation time and final generation

In an IEC system, the time spent in evaluating designs is a significant factor in reducing the evaluation load of the users.

Figure 12 shows the average and standard deviations of the time spent in evaluating the designs for each system. The evaluation time is measured from the point of initial presentation of the designs until finishing the evaluations are finished in the final generation. The evaluation times required for E-IEC 1 and 2 systems were shorter than those required for the NIEC system because subjects evaluate the designs manually in NIEC. In the E-IEC system, they only view the presented designs or their favorite designs. Moreover, the evaluation time required for E-IEC 2 is shorter than that for E-IEC 1 because subjects using the E-IEC 2 system are viewing the designs freely.

The Friedman test results show that when the significance level is 5%, the statistically significant differences in the evaluation times of the system are confirmed. Moreover, Scheffe’s multiple comparison results show that when the significance level is 5%, the evaluation times for E-IEC 1 and 2 are significantly shorter than that of NIEC.

Figure 13 shows the distribution of the evaluation times. Almost all subjects finish the evaluation of designs within 240 [s] in the E-IEC 1 and 2 systems. However, approximately only 70% of all subjects spend within 240 [s] in the evaluating the designs in NIEC.

We also investigate the results of the evaluation time for each generation. Figure 14 shows the results of the average evaluation times of each generation. All subjects spend longer time on design evaluation at the initial stage than after several evaluations in all systems because they evaluate designs carefully given that the systems present various designs at the initial stages. However, in E-IEC systems, the evaluation time of each generation is shorter than that of the NIEC system. Therefore, users can easily evaluate the E-IEC systems.

In addition, we survey the final generations of each system to confirm the differences in user evaluation iterations. Lower generations and total evaluation times are significant factors in reducing evaluation loads and generating designs that satisfy users.

Figure 15 shows the final generation results where all the systems are almost equal from 6 to 7. We perform the Friedman test to confirm the statistically significant differences in the final generation of each system. The Friedman test shows that when the significance level is 5%, the statistically significant differences in the final generation of each system are not confirmed. Therefore, the subjects repeat the evaluation of the designs in the same manner for all systems. However, the subjects only view the presented designs for evaluation in E-IEC systems.

![Figure 12: Results of evaluation time](image)

![Figure 13: Results of the distribution of evaluation time](image)

![Figure 14: Results of the evaluation time of each generation](image)

![Figure 15: Results of the final generations](image)
From all results, we confirm that the E-IEC system can not only generate various designs that satisfy users but also reduce their evaluation loads. Moreover, the evaluation times using the E-IEC system are significantly shorter than that of NIEC even if the number of generations among all systems is the same. Therefore, the E-IEC system is effective for creating an easy-to-use and effective IEC system.

Finally, we find the following issue for improvement the E-IEC system.

- Accuracy improvement of user gaze sampling
  The gaze information of the E-IEC system includes some noise data because user may only view just to be interested in the designs. Therefore, we need to improve a method of gaze sampling as describe in 4.1.

5. CONCLUSION

We presented an IEC system that applied user gaze information for evaluating candidate solutions. We performed the evaluation experiments to verify the effectiveness of the E-IEC system. The experimental results showed that the E-IEC system can significantly reduce the evaluation loads of users compared to the NIEC system and can generate designs that satisfy users. In future study, we will create gaze-based IEC application based on the system outlined in the present paper. We also need to improve accuracy of gaze sampling for removing noise of gaze sampling.

ACKNOWLEDGEMENTS

This research was partially supported by MEXT. KAKENHI (16K16141).

REFERENCES

1. Takagi, H.; Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation, Proceedings of the IEEE, 89(9), pp.1275-1296, 2001.
2. Lai, C. C., and Chen, Y. C.; A user-oriented image retrieval system based on interactive genetic algorithm, IEEE Transactions on Instrumentation and Measurement, 60(10), pp.3318-3325, 2011.
3. Marques, V. M., Reis, C., and Tenreiro Machado, J. A.; Interactive evolutionary computation in music, 2010 IEEE International Conference on Systems Man and Cybernetics (SMC), Istanbul, pp.3501-3507, 2010.
4. Ishibashi, K., and Miyata, K.; Statistics-based interactive evolutionary computation for color scheme search, International Journal of Affective Engineering, 14(1), pp.31-41, 2015.
5. Akase, R., and Okada, Y.; Web-based multiuser 3D room layout system using interactive evolutionary computation with conjoint analysis, Proceedings of the 7th International Symposium on Visual Information Communication and Interaction (VINCI ’14), Sydney, pp.178-187, 2014.
6. Yoon, D., and Kim, K. J.; 3D Game model and texture generation using interactive genetic algorithm, Computers in Entertainment, 14(1), pp.1-16, 2017.
7. Takenouchi, H., Tokumaru, M., and Muranaka, N.; Tournament-style evaluation using kansei evaluation, International Journal of Affective Engineering, 12(3), pp.395-407, 2013.
8. Takenouchi, H., Tokumaru, M., and Muranaka, N.; Interactive evolutionary computation using a Tabu Search algorithm, IEICE Transactions on Information and System, E96-D(3), pp.673-680, 2013.
9. Unehara, M., Yamada, K., and Shimada, T.; Subjective evaluation of music with brain wave analysis for interactive music composition by IEC, Joint 7th International Conference on Soft Computing and Intelligent Systems and 15th International Symposium on Advanced Intelligent Systems (SCIS&ISIS 2014), Kitakyusyu, pp.66-70, 2014.
10. Fukumoto, M., Ogawa, S., Nakashima, S., and Imai, J.; An extended interactive evolutionary computation using heart rate variability as fitness value for composing music chord progression, Journal of Advanced Computational Intelligence and Intelligent Informatics, 15(9), pp.1329-1336, 2011.
11. Moniri, M. M., Valcarcel, F. A. E., Merkel, D., and Sonntag, D.; Human gaze and focus-of-attention in dual reality human-robot collaboration, 2016 12th International Conference on Intelligent Environments, London, pp.238-241, 2016.
12. Pallez, D., Collard, P., Baccino, T., and Dumercy, L.; Eye-tracking evolutionary algorithm to minimize user’s fatigue in IEC applied to interactive one-max problem, GECCO ’07 Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation, London, pp.2883-2886, 2007.
13. Holmes, T., and Zanker, J.; Eye on the prize: using overt visual attention to drive fitness for interactive evolutionary computation, GECCO ’08 Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation, Atlanta, pp.1531-1538, 2008.
14. García-Saravia, J., Salas-Morera, L., García-Hernández, L., and Cabrera, A. A.; Application of an eye tracker over
facility layout problem to minimize user fatigue, International Work-Conference on Artificial Neural Networks (IWANN 2017), pp.145-156, 2017.

15. Jung, J., Matsuba, Y., Mallipeddi, R., Funaya, H., Ikeda, K., and Lee, M.; Evolutionary programming based recommendation system for online shopping, 2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Kaohsiung, pp.1-4, 2013.

16. Takenouchi, H., and Tokumaru, M.; Performance Evaluation of interactive evolutionary computation applying gaze information, The 4th International Symposium on Affective Science and Engineering and the 29th Modern Artificial Intelligence and Cognitive Science (ISASE-MAICS 2018), Spokane, A2-3, 2018.

Hiroshi TAKENOUCHI (Member)
Hiroshi Takenouchi is an Assistant Professor of Faculty of Information Engineering at Fukuoka Institute of Technology, Japan where he has worked since 2014. He received B.E., M.E. and D.E. degrees in Engineering from Kansai University in 2008, 2010 and 2013, respectively. In 2013, he had been a part-time researcher and postdoctoral fellow of Kansai University. His current research interests include fuzzy reasoning and an application with Interactive Evolutionary Computation. He is a member of Japan Society for Fuzzy Theory and Intelligent Informatics, Japan Society of Kansei Engineering, Institute of Electronics, Information and Communication Engineers and IEEE.

Masataka TOKUMARU (Member)
Masataka Tokumaru is a Professor of Faculty of Engineering Science at Kansai University in Osaka, Japan since 2016. He received the B.E. and M.E. degrees in Engineering from Kansai University in 1994 and 1997. During 1997-2007, he had been a Research Associate or a Lecturer of Department of Electronics at Kansai University. He received Dr. (Eng.) degree from Osaka City University, Japan in 2001. His research interests are affective computing and evolutionary computation. He is a member of Japan Society of Kansei Engineering, Institute of Electronics, Information and Communication Engineers, Japan Society for Fuzzy Theory and Intelligent Informatics, Human Interface Society, Japanese Society for Artificial Intelligence and IEEE.