Discussion of Offspring Generation Method for interactive Genetic Algorithms with Consideration of Multimodal Preference

Fuyuko Ito[1], Tomoyuki Hiroyasu[2], Mitsunori Miki[3], and Hisatake Yokouchi[2]

[1] Graduate School of Engineering, Doshisha University,
1-3 Tatara Miyakodani Kyotanabe, Kyoto, Japan,
fuyuko@mikilab.doshisha.ac.jp
[2] Department of Life and Medical Sciences, Doshisha University
tomo@is.doshisha.ac.jp,hyokouch@mail.doshisha.ac.jp
[3] Department of Science and Engineering, Doshisha University
mmiki@mail.doshisha.ac.jp

Abstract. The interactive genetic algorithm(iGA) is a method to obtain and predict a user 's preference based on subjective evaluation of users, and it has been applied to many unimodal problems, such as designing clothes or fitting of hearing aids. On the other hand, we are interested in applying iGA to user 's preferences, which can be described as a multimodal problem with equivalent fitness values at the peaks. For example, when iGA is applied to product recommendation on shopping sites, users have several types of preference trends at the same time in product selection. Hence, reflecting all the trends in product presentation leads to increased sales and consumer satisfaction. In this paper, we proposed a new offspring generation method that enables efficient search even with multimodal user preferences by introducing clustering of selected individuals and generating offspring from each cluster. Furthermore, we performed a subjective experiment using an experimental iGA system for product recommendation to verify the efficiency of the proposed method. The results confirmed that the proposed method enabled offspring generation with consideration of multimodal preferences, and there was no negative influence on the performance of preference prediction by iGA.

1 Introduction

The interactive genetic algorithm(iGA)[1] is an optimization method in which users evaluate the solutions instead of the fitness function of a genetic algorithm(GA)[2] and it optimizes targets that could not be formulated by iterating the following steps: presentation of individuals to users, evaluation of individuals by users, selection, crossover, and mutation. However, there are a number of problems in iGA, such as prediction of fitness values, combating user fatigue[3], etc. In this study, we focused on prediction of fitness model or user 's preference and how to make progress in presenting population to users.

As iGA replaces the fitness function with a user 's preference, the landscape of preference differs widely depending on the target problems of iGA. If the target problem has only a unique solution, the preference landscape is unimodal. On the other hand, if there are some trends in a user 's preference, the landscape is multimodal. However, conventional iGA searches for only one of these trends. Toward the end of such a search, the population converges to an optimal solution and very few individuals that are similar to other optimal solutions are presented. Particularly, for multimodal function with equivalent fitness values at the peaks, it is preferable to present all peaks to users during the search.

Applying iGA to product recommendation on shopping sites is one such problem. Product recommendation, such as collaborative filtering(CF)[4, 5], is a technique utilized to present products that are likely to be bought based on a user 's preferences. When users select products on shopping sites, there are several types of preference trends at the same time, which can be described as multimodal functions with equivalent fitness values at the peaks.

Here, we discuss an offspring generation method of iGA that enables efficient search even if user 's preference is described as a multimodal function with equivalent fitness values at the peaks. The proposed method introduces population clustering to an iGA. The following section presents an application of iGA for product recommendation and associated problems. Then, the proposed offspring generation method considering multimodal preferences is described. Finally, the effectiveness of the proposed method was verified by experiment with a recommendation system using iGA.

2 Interactive Genetic Algorithms and Multimodal Preference

2.1 Interactive Genetic Algorithms

The interactive genetic algorithm (iGA)[1] is an optimization method based on subjective evaluation of users and the genetic operations of GA[2]. Subjective preference replaces the fitness function for evaluation operation of GA in iGA. Therefore, iGA has been applied as a method of sensitivity analysis to many problems that are difficult to formulate, such as the design of clothes and music composition. Aoki et al.[6] also confirmed that iGA is especially effective for users inexperienced with the target problem. IGA consists of the following steps:

1. Generate an initial population of N_{pop} individuals.
2. Present the individuals to the user.
3. Evaluate each individual based on the user's preference.
4. Select individuals to be kept for the next generation from the user's evaluation.
5. Exchange chromosomes between individuals and generate offspring.
6. Randomly change chromosomes to maintain diversity.
7. Terminate if the user obtains desired individual, else iterate steps 2 to 6.

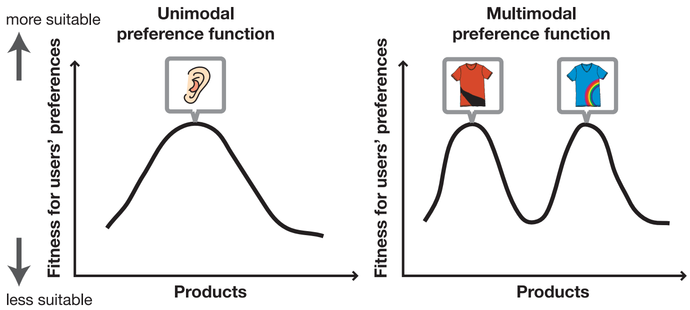


Fig. 1. Unimodal and multimodal preference functions

On the other hand, a number of daunting challenges of iGA remain: discrete fitness value input method, prediction of fitness values, interface for dynamic tasks, acceleration of iGA convergence, combination of evolutionary and non-evolutionary computation, active intervention, and theoretical research. In this study, we focused on prediction of fitness model of user 's preference and examined an offspring generation method to make progress on presenting population to users with multimodal preferences.

2.2 IGA for Multimodal Preference

When iGA is applied to actual problems, the landscape of the fitness function, the user preferences in this case, is very important. For example, in optimization of the parameter settings of a hearing aid, there exists the best combination of parameters that will help the user to hear. Thus, preferences regarding optimization of the hearing aid are described as a unimodal function, as shown in Fig.1, and the population converges on a unique optimal region. On the other hand, when users select products, such as clothes on shopping sites, each user generally has multiple preferences. Then, the preferences on product selection are described as a multi-peak function, as shown in Fig.1, and individuals that suit all preferences should be presented.

However, iGA is a method that searches for a solution in a unique optimal region[1] and it is not beneficial for users if all individuals converged to only a single optimal region of multimodal preference when the user 's preferences are multimodal. Therefore, an offspring generation method that responds to not only unimodal preferences, but also multi-peak preferences is necessary. To achieve this, a new offspring method using clustering is proposed in this paper.

3 Offspring Generation Method with Consideration of Multimodal Preference

3.1 Overview of the proposed method and its execution timing

To enable offspring generation considering multimodal preferences, each peak of preference function must be discovered. To obtain local optima of a multimodal function, Hocaoglu and Sanderson[7] introduced clustering to GA, and the population is divided into several sub-populations. These sub-populations evolve independently, exchanging individuals, and are merged at a certain interval. Moreover, sharing (or niching) reveals the distribution of local optima and maintains the diversity of the population by updating the fitness value of each individual based on degree of congestion. According to the former, we assumed that each peak of multimodal preferences corresponds to scattered regions in the design domain and proposed an offspring generation method taking multimodality into consideration using clustering of highly evaluated individuals.

In the proposed method, a user clicks favorite individuals and each individual is evaluated with regard to whether the user has clicked it or not. The regions that suit the user 's preference are specified from the design space by clustering, and offspring are generated from these regions. The number of clusters must be determined automatically and the number of regions is not known initially. However, most clustering methods, such as K-means or agglomerative clustering, require *apriori* knowledge of the number of clusters. Hence, multiobjective clustering with automatic k-determination (MOCK)[8], which automatically determines the number of clusters, was adopted in the proposed method. Moreover, the proposed offspring generation method is applied once at the mth generation of iGA, as shown in Fig.2. In other generations, conventional crossover and mutation operators are utilized to generate offspring. Thus, it is possible to generate offspring that are suitable to the user 's preference beyond specified regions.

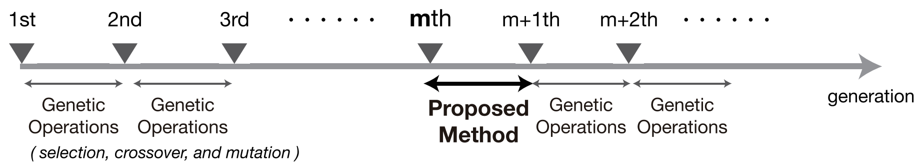


Fig. 2. Application timing of the proposed method for iGA

3.2 Procedure of Offspring Generation with Consideration of Multimodal Preference

Figure 3 shows an overview of the proposed method. In an n-dimensional design space, the proposed method generates N_{off} offspring as follows:

1. At the mth generation, all individuals selected by the users up to mth generation are divided into clusters, $\mathbf{C}_1$, ..., $\mathbf{C}_k$, ..., $\mathbf{C}_{N_{cluster}}$ in the design space, with the number of clusters $N_{cluster}$ is determined automatically in MOCK.
2. The offspring generation range R_k corresponding to cluster $\mathbf{C}_k$ is determined by the distribution range of individuals $\mathbf{I}_{k1}$, $\mathbf{I}_{k2}$, ..., $\mathbf{I}_{kl}$ in $\mathbf{C}_k$.
3. N_{off} offspring (as many as population size N_{pop}) are generated randomly from offspring generation ranges R_0, ..., $R_{N_{cluster}}$. Each range generates the same number of offspring.

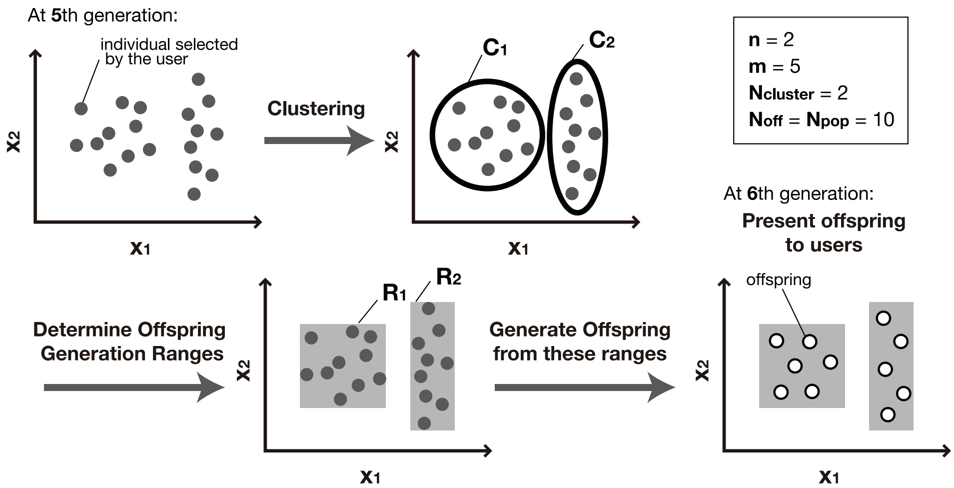


Fig. 3. Offspring generation procedure in the proposed method

4. Generated offspring are presented at the $m+1$th generation and operations of conventional iGA are iterated subsequently.

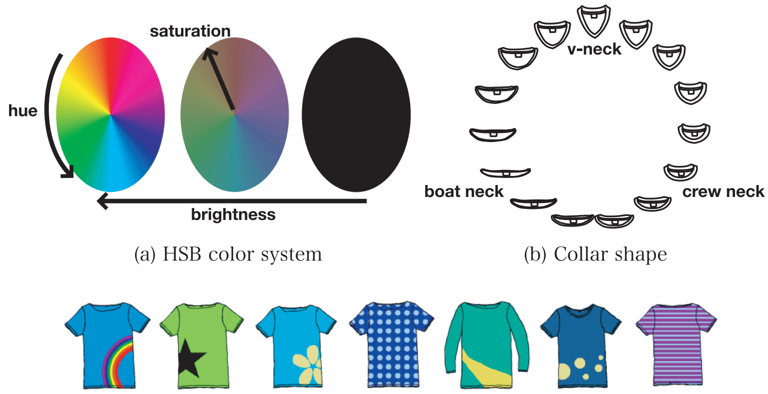
As described above, the proposed method specifies the regions that suit the user 's preferences from the distribution range of the individuals selected by the user up to the mth generation. These individuals are divided appropriately into clusters by MOCK, which determines the number of clusters, and it enables generation of offspring with consideration of not only unimodal but also multimodal preferences.

4 Assessment of Offspring Generation with Consideration of Multimodal Preference

4.1 Experimental iGA System for Product Recommendation

For assessment of the proposed offspring generation method, we developed a product recommendation system using iGA and adopted t-shirts as the target product. The users evaluate the presented products by clicking based on whether they are suitable for their preferences. Design variables of a t-shirt and the flow of the experimental system are shown below:

Design Variables The design variables of a t-shirt consist of color, shape, and pattern. Expressions of each design variable are shown in Fig.4. The color of a t-shirt is expressed by the HSB color system, which is similar to the human sense of color recognition. The HSB color system expresses color as a combination of hue (0 to 360 degrees), saturation (0 to 100), and brightness (0 to 100) as shown in Fig.4(a). These design variables of the HSB color system are continuous values. The shape of a t-shirt is described by the shape of the collar and the length



(a) HSB color system (b) Collar shape

(c) Patterns

Fig. 4. Design variables of a t-shirt

of the sleeves as discrete values. Fifteen collar shapes are constructed based on three basic shapes—boat neck, v-neck, and crew neck—as shown in Fig.4(b). The sleeve types are short sleeves and long sleeves. Eight t-shirt patterns including a solid color are shown in Fig.4(c), and the color of each pattern is determined from 9 colors: white, yellow-green, sky blue, blue, purple, pink, red, yellow, and black. The design variables of t-shirt patterns and their colors are discrete values.

Flow of Experimental System

- Generation and presentation of the initial population
  First, the system generates N_{pop} individuals as an initial population. The design variables of the initial population are determined randomly. Saturation and brightness are determined from the range of 75 to 100 to present a wide range of hue clearly.
- Evaluation and selection
  Users click the products on the screen that suit their preferences. Presented individuals are evaluated by whether they are clicked or not. Individuals clicked by the user are selected as parents. The number of parents is set to half the population size N_{pop}, and users are allowed to click up to half of N_{pop} individuals. If clicked individuals are less than the number of parents, the system automatically selects the remaining parent individuals from the population in increasing order of Euclidean distance from the clicked individuals.
- Crossover
  The crossover operator generates two offspring from two parents. Hue and collar shape of the offspring are determined from the acute angle made by two parents in the circular design space. Saturation and brightness of offspring are determined between the design variables of two parents in the design space. Other design variables of offspring are inherited from the parents.

- Mutation
  The mutation operator changes the design variables of offspring at random based on mutation rate.

4.2 Experimental Overview

An experiment was performed to determine whether the proposed method generates offspring with consideration of multimodal preferences. In this experiment, we adopted the experimental product recommendation system described in Section 4.1. Subjects utilized the system with the proposed offspring generation method as the proposed system and the system with only the conventional genetic operations as the conventional system. Each system showed 20 individuals in one generation and presented t-shirts in the interface, as shown in Fig.5. The subjects consisted of 14 men and 6 women ranging in age from 21 to 27 years, and the order of the systems used was counterbalanced among subjects. They were instructed to select their favorite t-shirts in each system and to evaluate the t-shirts presented in each system for 10 generations. The crossover rate and mutation rate were set to 1.0 and 0.2, respectively. The proposed system generated offspring using clustering after evaluation of the 5th generation and presented the generated individuals at the 6th generation. Hue and collar shape were adopted as the target design variables for clustering, and the number of clusters was set in the range of 1 to 3 in this experiment. Subjects answered questionnaires after experimenting with each system.

4.3 Results and Discussion

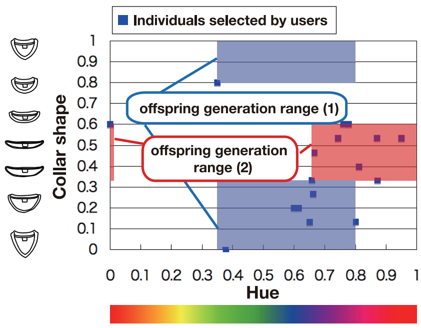
First, we discuss whether the proposed method generated offspring that considered multimodal preferences or not. Figure 6(a) shows the offspring generation ranges of subject A determined by the proposed method and the individuals selected by the user up to the 5th generation before applying the proposed method. The horizontal and vertical axes show hue and collar shape, respectively. The offspring generation range (1) seemed to be divided horizontally, and range (2) also seemed to be divided vertically. However, they were not divided because

Fig. 5. The experimental system interface

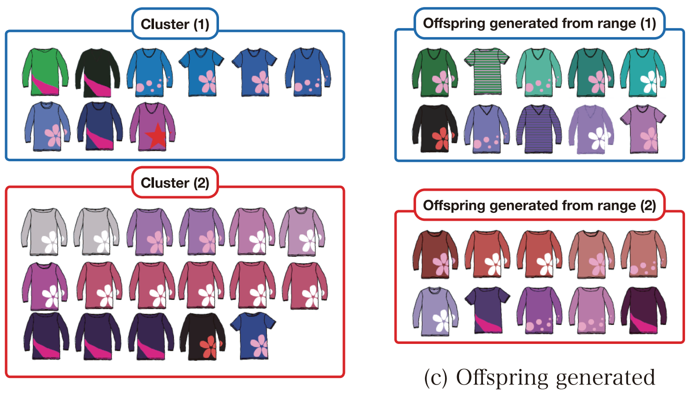
the design space was circular. The two ranges overlapped each other around the range of purple, but selected individuals were divided into two clusters (see Fig.6).

Next, Fig.6(b) shows the results of clustering of selected individuals. Clusters (1) and (2) consisted of cool and warm color t-shirts, respectively. Therefore, subject A had multimodal preferences and liked both cool and warm colors. Cluster (1) also consisted of v-neck t-shirts, while cluster (2) consisted of boat neck t-shirts. Black and flower-patterned t-shirts were included in both clusters because brightness and patterns of t-shirts were not adopted in clustering in this experiment. Thus, target design variables for clustering should be discussed in future studies.

Moreover, Fig.6(c) shows the list of individuals presented at the 6th generation that were generated from these offspring generation ranges. Purple t-shirts were generated from both ranges because they overlapped with each other (see Fig.6(a)). However, the population did not converge on one of the user 's multiple preferences. These trends of experimental results were similar to 12 of 20 subjects



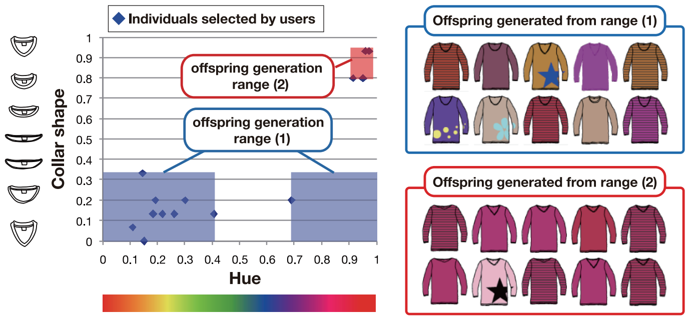
(a) Selected individuals and offspring generation ranges



(b) Clustering result

(c) Offspring generated by the proposed method

Fig. 6. Example of offspring generation considering multimodal preferences



(a) Selected individuals and offspring generation ranges

(b) Offspring generated by the proposed method

Fig. 7. Example of failure by the proposed method

and thus the proposed method would be able to generate offspring considering multimodal preferences.

In addition, a case where the proposed method did not perform efficiently is also discussed. Figure 7(a) shows the offspring generation ranges and the individuals that were selected by subject B up to the 5th generation. Figure 7(b) shows the offspring generated from these ranges. Subject B selected many green t-shirts by the 5th generation (see Fig.7(a)) and one of the green t-shirts was selected five times. However, no green t-shirts were included in the offspring generated by the proposed method (see Fig.7(b)). There was a possibility that green t-shirts would be generated from the offspring generation range (1), but it was not because the offspring were generated without considering the distribution of selected individuals in each range. To achieve this, offspring generation with a probabilistic model should be discussed in future studies.

A questionnaire study was used to assess whether the proposed offspring generation method had negative effects on the performance of preference learning of iGA. The question was, " Which system presented more favorable t-shirts? " The subjects chose their responses from among 4 choices as shown in Fig.8. Eighty-five percent of the subjects responded that the proposed system or both systems presented individuals that suited their preferences (see Fig.8). Thus,

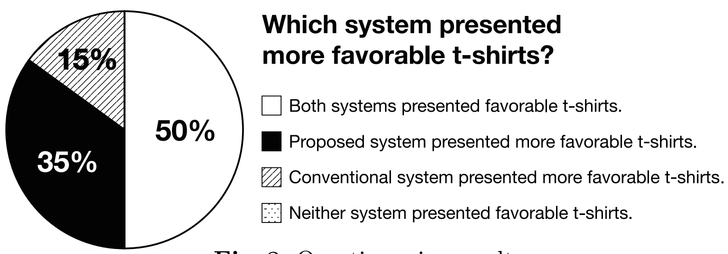


Fig. 8. Questionnaire results

subjective evaluation confirmed that the proposed offspring generation method using clustering did not negatively affect the performance of preference learning of the iGA.

5 Conclusions and Future Works

The iGA is an effective method to introduce a user 's preference to a system. It is capable of converging on unique optima by conventional iGA when a user 's preference is unimodal. On the other hand, a mechanism that enables efficient search even if a user 's preference is multimodal with equivalent fitness values at the peaks is needed. Therefore, we proposed a new offspring generation method considering multimodal preferences. The proposed method introduces clustering of selected individuals to obtain local-optima and generates offspring from each cluster. We performed a subjective experiment using an experimental iGA system for product recommendation. The experimental results indicated that the proposed method could appropriately determine trends in a user 's preference and generate offspring with consideration of multimodal preferences. In future work, we will discuss the target design variables for clustering and offspring generation with a probabilistic model from specified regions considering dependencies among design variables. Furthermore, we will also improve the method of specification of regions that suit a user 's preference.

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. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Professional (1989)
3. Llorà, X., Sastry, K., Goldberg, D.E., Gupta, A., Lakshmi, L.: Combating user fatigue in igas: partial ordering, support vector machines, and synthetic fitness. In: GECCO '05: Proceedings of Genetic Evolutionary Computation Conference pp.1363–1370 (2005)
4. Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R., Riedl, J.: Grouplens: applying collaborative filtering to usenet news. Commun. ACM **40**(3) pp.77–87 (1997)
5. Sarwar, B., Karypis, G., Konstan, J., Reidl, J.: Item-based collaborative filtering recommendation algorithms. In: WWW '01: Proceedings of the 10th international conference on World Wide Web, pp.285–295 (2001)
6. Aoki, K., Takagi, H.: Interactive GA-based design support system for lighting design in 3-D computer graphics. The transactions of the Institute of Electronics, Information and Communication Engineers **81**(7) pp.1601–1608 (1998)
7. Hocaoglu, C., Sanderson, A.C.: Multimodal function optimization using minimal representation size clustering and its application to planning multipaths. Evolutionary Computation **5**(1) pp.81–104 (1997)
8. Handl, J., Knowles, J.: Multiobjective clustering with automatic determination of the number of clusters. Technical Report TR-COMPSYSBIO-2004-02, UMIST, Department of Chemistry (2004)