Visualization of Search Process and Improvement of Search Performance in Multi-Objective Genetic Algorithm

Daisuke Yamashiro, Tomohiro Yoshikawa, *Member, IEEE*, and Takeshi Furuhashi, *Member, IEEE*

Abstract— Performance in searching solutions by Multi-Objective Genetic Algorithm (MOGA) depends on genetic operators and/or their parameters. For comparison of the performance with some genetic operators and/or parameters, it has been usually employed the transitions of fitness values through actual applications or the number/performance of acquired Pareto solutions in multi-optimization problems. This paper proposes a visualizing method of search process for MOGA, which can visualize relative distances among chromosomes in search process and give information of not only the performance but also the effects of the genetic operations such as the diversity of chromosomes. This method uses Self-Organizing Map (SOM) for the visualization. This paper applies Non Dominated Sorting Genetic Algorithm-II (NSGA-II) to ZDT2 and FON test functions and shows obtained nondominated solutions and visualization results. This paper also shows that the visualized data enables us to interpret the differences in search processes and to get new information to determine efficient genetic operators and their parameters.

I. INTRODUCTION

Recently, the research of Multi-Objective Genetic Algorithm (MOGA) [1], [2], [3] has been emphasized and reported. Genetic Algorithm (GA) is one of the most effective methods for searching nondominated solutions within practical time in multi-objective optimization problem. Searching solutions in MOGA progresses with genetic operators, which are based on biological evolutional process such as selection-reproduction, crossover, mutation and so on. The search performance of MOGA drastically depends on the genetic operators and their parameters, e.g. number of individuals, probability of crossover, that of mutation. The comparison of their performance among various kinds of operators or parameters by trial and error with actual applications has been the only way to acquire the effects of them and determine the best one so far. If the objective problem is simple and the solution space is less or equal to 3 dimensions, we can display it in graphic and visually see the search process. It is, however, difficult to visualize it in actual combinational optimization problems that are effective to apply GA because of their multidimensional searching space. One of the actual multi-objective optimization problems is Nurse Scheduling

Problem (NSP). To cite an application of MOGA to NSP with 20 nurses, 30 days and 7 working patterns, if each working pattern is defined as independence, the searching space is 4,200 dimensional one. Then the genetic operators and parameters have been determined by their comparison described above through the transition of fitness values. In multi-optimization problem such as NSP, the fitness values and the number of acquired nondominated solutions are also the indexes of the performance for them. However, they give us just the information of their performance, and we can not grasp the search process in this 4,200 dimensional space.

Visualization is one of the effective methods to grasp search process of GA. In Interactive Evolutionary Computing (IEC), Takagi et al. [4] have proposed the acceleration method for evolutional convergence through visualized landscape of multidimensional solution space. It shows the visualized space to the user, and he/she selects an area gathering high fitness values, then it generates new individuals corresponding to this area in the next generation. This method is similar to the proposed method in this paper in terms of using Self Organizing Map (SOM) [5], [6], [7] for the visualization and using the information of chromosomes as the input vectors for SOM. It is, however, different in visualizing the landscape of whole searching space and the objective of visualization that is the acceleration of convergence by predicting the search area with high fitness values. Additionally, this method makes positional information of chromosomes including fitness value mapped through 2-dimensional space. However, this method has some problems. For example, it is difficult to apply this method to multi-objective problems which have a number of fitness axes and to verify if the landscapes in multi-dimensional space are preserved on mapped 2-dimensonal space.

In this study, we aim to get new information and proper index to determine adaptive genetic operators and their parameters through visualization of search process in MOGA.

This paper proposes a visualization method of relative distances among individuals in MOGA to grasp the search process easily. We focus on that it is much easier to visualize relative distances among individuals though it is difficult to do multidimensional searching space itself. The proposed method enables us to understand the appearance of the convergence or diversity of individuals, the collecting and distributing of new individuals generated by crossover and mutation, which the transition of fitness value does not have. The proposed method can also show the generating process of new individuals by genetic operators in addition to relative relationship among individuals. That leads us to visually get

Daisuke Yamashiro is with the Department of Computational Science and Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, JAPAN (phone: +81-52-789-2793; fax: +81-52-789-3166; email: daisuke@cmplx.cse.nagoya-u.ac.jp).

Tomohiro Yoshikawa is with the Associate professor of Department of Computational Science and Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, JAPAN (email: tom@fcs.coe.nagoya-u.ac.jp).

Takeshi Furuhashi is with the Professor of Department of Computational Science and Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, JAPAN (email: furuhashi@cse.nagoya-u.ac.jp).

new information to set genetic operators and their parameters in a few trials.

The proposed method is independent of the number of the objective functions and easily applied to multi-objective problems because of visualizing relative location among chromosomes. There is no assurance that the proposed method can preserve the distance among chromosomes in multi-dimensional space onto the visualized space, however, we can verify the relationships of the distances as long as the distances among chromosomes are defined.

This paper investigates the effectiveness of the proposed method by applying it to multi-objective functions. One of the most efficient methods to search nondominated solutions for multi-objective problems is Non Dominated Sorting Genetic Algorithm-II (NSGA-II) [8], which is widely used for solving multi-objective and multi-dimensional problems. One of the main characteristics of NSGA-II is to keep the diversity of chromosomes. It stores and renews at each generation the acquired nondominated solutions into the Pareto population. The particular chromosomes with less density in the Pareto population are given the priority to be applied genetic operations. Each density of chromosome is determined by calculating the distances with the next solutions of both side along the Pareto front. That is, the chromosomes of nondominated solutions with less density have more chance to be evolved and to generate new chromosomes around them. It leads to keeping the diversity of chromosomes to generate more kinds of Pareto solutions. This paper applies the proposed method to NSGA-II for ZDT2 and FON test functions [9] and investigates the effectiveness of the genetic operators in NSGA-II for each test function by not only transition of Pareto solutions but also results of visualization.

This paper shows that the proposed method enables us to clarify the difference of the features of search process in NSGA-II for each test function and the effect of the genetic operators. It investigates the reason of superior performance through the visualized results of the divergence of individuals and the generated process by the operators. This paper also shows that the designers can feed back the acquired features into the genetic operators and the parameters.

This paper is organized as follows: Chapter 2 describes the proposed method, chapter 3 shows the experimental results and investigation, and this paper concludes in chapter 4.

II. PROPESED METHOD

A. Flow of the Proposed Method

Figure 1 shows the overview of the proposed method. The proposed method defines relative distances among individuals in GA and visualizes them. Then, we can understand the search process visually and feed back its information into efficient or demanded genetic operators and their parameters. Though the information of the genotypes of individuals is employed to determine the relative distances among them in this paper, this method also covers that of the fitness values in multi-objective problems such as NSP. When we employ the fitness values to determine the distances, each
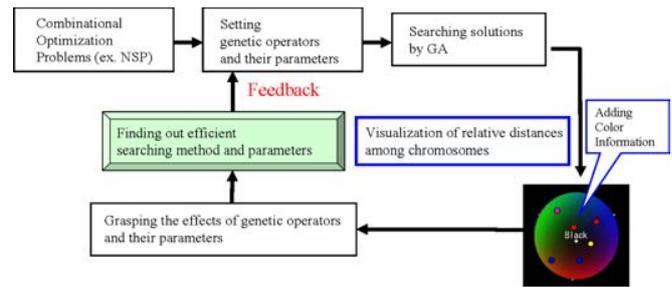

Fig. 1. Feedback into genetic operators from visualization of search process

element of a vector is represented by each objective fitness function's value, and we can understand the search process on the fitness space. This paper defines Euclidean distance as the relative distance. We can also define proper distances such as Hamming distance, Manhattan distance and so on according to each problem or coding.

The flow of the proposed method is as follows:

(Step 1) Definition of the relative distance among chromosomes suitable to each problem or objective regarding the chromosomes as multidimensional data.

(Step 2) Setting the employed genetic operators and their parameters.

(Step 3) Input the chromosomes into SOM as input vectors in each generation.

(Step 4) The population is mapped onto the visualized space with color information (nondominated solutions, generated type by crossover or mutation) added to individuals.

(Step 5) Understanding and investigating the features of employed genetic operators and their parameters through visualized search process.

(Step 6) Feedback of the acquired information from (Step5) into (Step2).

B. Visualization by SOM

This paper employs SOM to visualize relative relationship among individuals in MOGA. SOM is one of the unsupervised learning algorithm, and it can map relative relationships among data on 2 dimensional space. SOM consists of two layer. One is called "Input Layer" and the other is called "Map Layer". In SOM, neurons in the map layer (output nodes) are generally assigned in grid pattern, and the number of neurons in the input layer (input nodes) is equal to that of input data's dimension. Each output node is connected to all input node with each weight vector. Then the number of weight vectors's dimension is equal to that of inputs. The main characterisitc of SOM is that the similar vectors among the input data are put closer on the map nodes. Especially, this paper employs Spherical SOM [10], [11], [12] to keep higher phase relation than normal SOM.

Mode of expression on SOM is explained briefly. Spherical SOM is able to rotate around and it is possible to be seen from any direction of horizontal/vertical 360-degree. In this paper, both figures of a certain direction and the opposite side

rotated 180-degree are shown when it shows the visualization result. Each individual is plotted on the spherical SOM with small square shape. Some color information is added to these plotted points in actual figure. Green means that the individual is nondominated solution without genetic operators. Yellow means that the individual was generated by crossover and it became nondominated solutions at that generation, blue was by mutation. Red means that the individual was generated by crossover and it was selected in that generation, purple was by mutation. If green and another color are plotted on the same point, the other color is displayed over the green. In the proposed method, the designer can add any colors onto the plotted points according to the required information. This paper employed the above colors to visualize the generated process of new individuals by genetic operators with the information if they had become nondominated ones or not. The more kinds of colors are employed, the more information can be shown. However, too many colors make the visualization result difficult to grasp.

The dark parts on the Spherical SOM represent gaps of the distance in input vectors, that is, the individuals surrounding the dark part are relatively far, and the light part means they are close. When some individuals are gathered in the area of light part, it represents a group consisted of similar chromosomes in that generation.

In this paper, the learning rate of SOM is set to 0.2 and the number of iterations is 10000, respectively. This paper employs nondominated solutions in the stored Pareto population and new individuals generated by crossover and mutation as input vectors. The input vectors from chromosomes are converted variables by eq. (2) or (4), respectively.

III. EXPERIMENT

This chapter shows the visualization results by the proposed method and discusses them with the obtained Pareto solutions and the generated process of individuals. NSGA-II was applied to ZDT2 and FON test functions. In this paper, 1-point crossover was employed and when a chromosome was chosen to mutation, randomly selected 1 gene was inverted. The genetic parameters were as follows: the population size for genetic operations was 100, the number of generations was 250, the probability of crossover was 0.9 and that of mutation was 0.1, respectively. These genetic parameters were employed referring to [8].

A. ZDT2 Test Function

This Section employed ZDT2 test functions for the objective function. The number of input variables of ZDT2 was 30 (n=30) in eq. (1). ZDT2 test function is shown in eq. (1).

$$minf_1(x) = x_1$$
$$minf_2(x) = g(x)[1 - (x_1/g(x))^2]$$
$$g(x) = 1 + 9(\sum_{i=2}^{n} x_i)/(n-1) \quad (1)$$
$$(0 \le x_i \le 1, n = 30)$$

The chromosomes were binary coded and the length was 300 bits. Each variable was represented by 10 bits and decoded to discrete value x' in closed interval [0, 1023]. Then x' was converted to corresponding value x by eq. (2).

$$x = \frac{x'}{2^{10} - 1} \quad (2)$$

1) Pareto Solutions: This subsection shows a result of obtained Pareto Solutions by NSGA-II. The obtained Pareto solutions are shown in Fig. 2. The horizontal axis represents
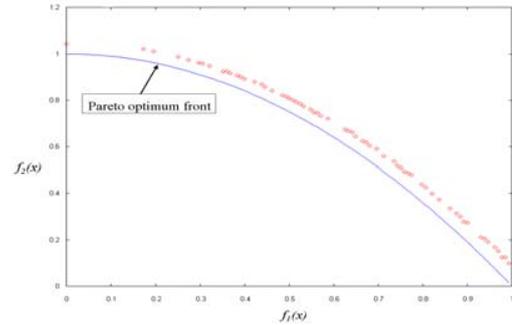

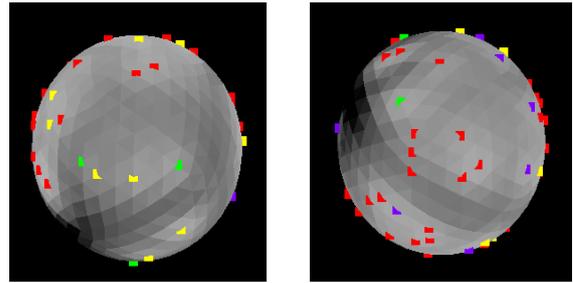Fig. 2.   Pareto solutions of ZDT2 by NSGA-II

Fig. 3.   Result of visualization at the 150th generation in ZDT2
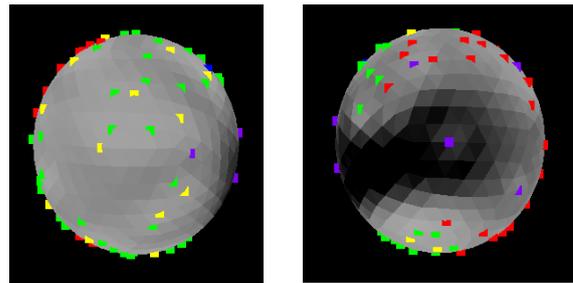

Fig. 4.   Result of visualization at the 250th generation in ZDT2

$f_1(x)$ and the vertical axis is $f_2(x)$. The small squares in Fig. 2 represent acquired Pareto solutions and the solid line represents Pareto optimum solutions of ZDT2. Figure 2 shows that a variety of Pareto solutions and continuous Pareto front were obtained, though they could not reach the optimum solutions.

2) Results of Visualization: Figure 3 and 4 show the visualization results at the 150th and 250th generation, respectively. Both figures show that the chromosomes were spread and covered the whole of spheres, which shows the

diversity of solutions in these generations. The effectiveness to keep chromosomes' diversity of NSGA-II could be confirmed by these visualization results.

In Fig. 3, the color information gives that some local parts have new nondominated solutions generated by genetic operations. It is thought that the operation of NSGA-II that gives solutions with less density more chance to generate new chromosomes worked well and they generated new nondominated solutions around them. In these generations, the evolution was done with high diversity.

In Fig. 4, the nondominated solutions are distributed evenly and the genetic operations generated new chromosomes around them. A lot of nondominated solutions are still generated especially by crossover. In these generations, genetic operations gave local search and it took a role to increase the kinds of Pareto solutions on the place between the parent chromosomes.

B. FON Test Function

This Section employed FON test function for the objective function. The number of input variables of FON was 4 (n=4) in eq. (3). FON test function is shown in eq. (3).

$$min f_1(x) = 1 - exp(-\sum_{i=1}^{n}(x_i - \frac{1}{\sqrt{n}})^2)$$

$$min f_2(x) = 1 - exp(-\sum_{i=1}^{n}(x_i + \frac{1}{\sqrt{n}})^2)$$

$$(-4.096 \leq x_i \leq 4.096, n = 4)$$

(3)

The chromosomes were binary coded and the length was 52 bits. Each variable was represented by 13 bits and decoded to discrete value x' in closed interval [0, 8191]. Then x' was converted to corresponding value x by eq. (4).

$$x = -4.096 + x'\frac{8.192}{2^{13} - 1}$$

(4)

1) *Pareto Solutions:* This subsection shows the results of obtained Pareto Solutions by NSGA-II and discusses them. The Pareto solutions are shown in Fig. 5 and 6. The horizontal and vertical axes represent $f_1(x)$ and $f_2(x)$, respectively. These figures show that more than half of the obtained Pareto solutions were variety and continuous, and they reached optimum solutions. However, the other Pareto solutions in the part of high f_1 value were poor in diversity and discontinuous, and they could not reach Pareto optimum solutions. In addition, the solutions at the 250th generation were not much improved from those at the 100th generation. It means that effective search had not progressed since the 100th generation.

It is difficult to find why it did not work well just from these figures. This paper applies the proposed method to this search in the next subsection and investigates the features of the generated process of individuals by genetic operators to clarify the search process.
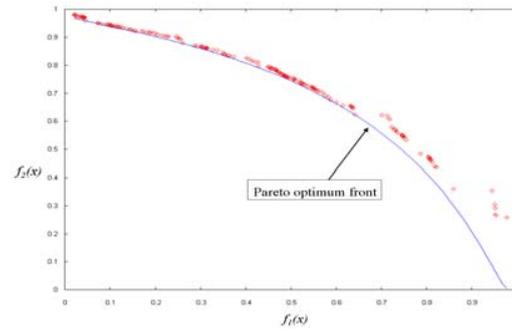

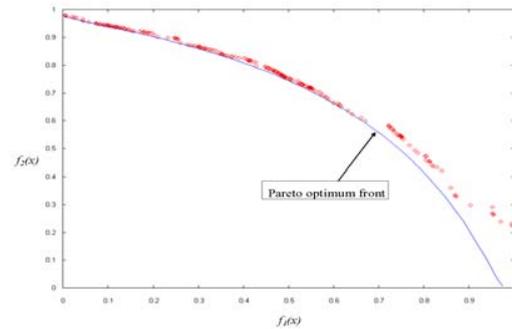
Fig. 5. Pareto solutions of FON at the 100th generation

Fig. 6. Pareto solutions of FON at the 250th generation

2) *Results of Visualization:* Figure 7, 8 and 9 show the visualization results at the 25th, 100th and 250th generation, respectively. At the 19th generation or later, the size of stored Pareto population was bigger than that for genetic operations. That is, new chromosomes were generated by genetic operations only from the stored nondominated solutions to which with less density more priority was given.

At the 25th generation, the features were similar to those in Fig. 3. NSGA-II generated some groups such as the circles in Fig. 7 keeping the evolution with high diversity. There were many yellow and blue squares in the groups, which meant new nondominated solutions were generated by crossover or mutation. In these generations, NSGA-II worked well to generate effective and more kinds of Pareto solutions.

At the 100th generation, most of the color of generated chromosomes were red. It means they were selected because they became inferior solutions. The conditions of the operations for this experiment, which are that the probability of crossover was much higher than that of mutation and the less density gave the priority to the chromosomes for genetic operations, give a following interpretation. In the part of high f_1 value in Fig. 5 and 6 of FON test function, the crossover with low density, i.e. isolated chromosomes did not generate new nondominated solutions around the chromosomes, then the chromosomes with low density kept isolated and being given the priority for genetic operations. It caused that the visualization result at the 250th generation in Fig. 9 showed almost no difference from that at the 100th generation in Fig. 8.
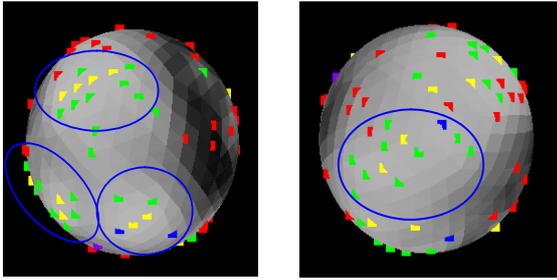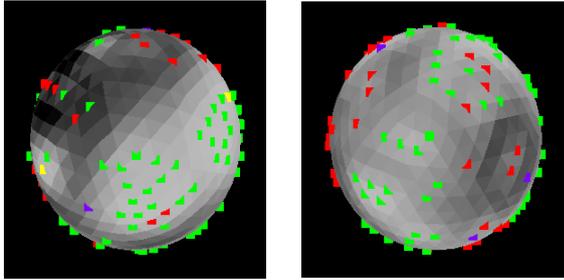
Fig. 7. Result of visualization at the 25th generation in FON

Fig. 8. Result of visualization at the 100th generation in FON

Fig. 9. Result of visualization at the 250th generation in FON

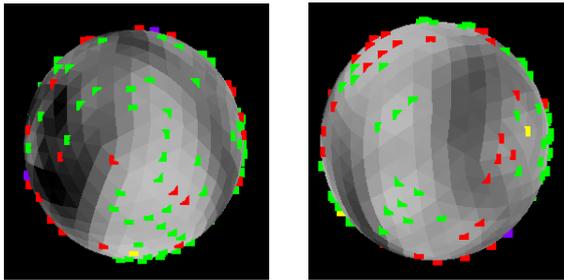Fig. 10. Pareto solutions of FON after feedback

Fig. 11. Result of visualization at the 250th generation in FON

It is also interesting results by the visualization. After the 100th generation, though not many new nondominated solutions were generated, the ratio of those by mutation was higher than by crossover. It is thought that the probability of mutation was too small in this experiment while the mutation took a role to generate new nondominated solutions around the chromosomes.

As the result of visualization from the 1st generation to the 250th generation of both test functions and the acquired Pareto solutions in Fig. 2, 5 and 6, this paper leads the features of NSGA-II in these test functions as follows:

(1) NSGA-II is strongly effective to keep the diversity of chromosomes.
(2) If crossover between relatively far chromosomes generates new chromosomes of nondominated solutions around them, the priority operation for low density solutions works well and many kinds of nondominated solutions are generated and distributed evenly along Pareto front.
(3) On the other hand, if the crossover generates inferior solutions, the priority operation prevents their improvement
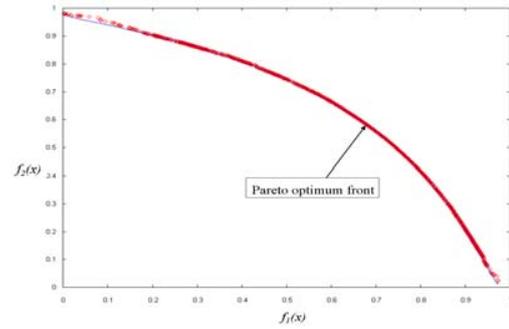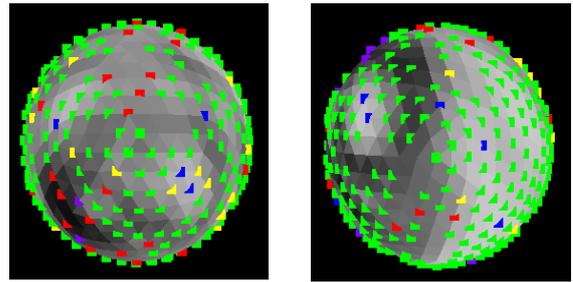
and evolution.
(4) The crossover takes a role to increase the kinds of nondominated solutions.
(5) In the case of (3), mutation takes an effective role to generate nondominated solutions instead of crossover.

This paper tries to feed back these features into the genetic operations in FON test function to improve the search performance in next subsection.

3) Feedback: The purpose of this study is to develop visualization method of search process in order to acquire the proper index to improve the genetic operators and reset genetic parameters. This subsection tries to feed back the features of search process which are acquired through visualization results. The feedback information to improve search performance is as follows:

(1) No priority for less density solutions, that is, the chromosomes for genetic operations are randomly chosen from the Pareto population.
(2) The probability of mutation is changed to 0.5 after the 100th generation.

The above feedback (1) and (2) are for the genetic operation and the genetic parameters, respectively. Figure 10 is the result of obtained Pareto solutions and Fig. 11 is that of visualization at the 250th generation. These figures show the feedback of this experiment was successful. Comparing with Fig. 6, much more kinds of Pareto solutions were acquired along the Pareto optimum front and they were evenly distributed in Fig. 10. Figure 11 also shows much improvement. The chromosomes with nondominated solutions are relatively distributed widely, and both crossover and mutation generated a lot of new nondominated solutions.

In this way, the proposed method enables us to get more and detail information in terms of search process and help to improve search performance.

IV. CONCLUSIONS

This paper proposed the visualization method of relative distances among individuals in MOGA to grasp the search process in terms of the effect on the genetic operators. The purpose of this research is to acquire the proper index to improve genetic operations and their parameters and feed back the acquired information through the visualization into them. This paper applied the proposed method to two multi-objective test functions by NSGA-II. It showed the proposed method enabled us to understand the convergence or diversity of individuals and collection or distribution of new individuals by adding the color information on SOM. The experimental results showed the visualized investigation enabled us to clarify the difference of the features in search process and effective feedback was done into the genetic operators and their parameters.

Future work is more investigation of the proposed method applying it to other genetic operators or parameters and real-world problems. Recently, the balance between diversity and convergence of individuals has been discussed actively [13]. The proposed method can be effective for this investigation. It is one of the application of this method to dynamical change of Evolutionary Strategy (ES) according to the search state. We will also apply this method to NSP that authors are carrying out [14] to find the features of efficient genetic operators for NSP. One of the applications of this method in NSP is to show plural candidates of scheduling tables using the relative distances among them.

ACKNOWLEDGMENT

REFERENCES

[1] H. Ishibuchi and T. Murata, "A multi-objective genetic local search algorithm and its application to flowshop scheduling," in *IEEE Trans. on Systems, Man, and Cybernetics - Part C: Applications and Reviews, Vol. 28, No. 3*, 1998, pp. 392–403.

[2] T. Murata, H. Nozawa, Y. Tsujimura, M. Gen, and H. Ishibuchi, "Effect of local search on the performance of cellular multi-objective genetic algorithms for designing fuzzy rule-based classification systems," in *Proc. of the 2002 Congress on Evolutionary Computation*, 2002, pp. 663–668.

[3] A. J, "Application of a multi objective genetic algorithm to engineering design problems," in *Proc. of the 2nd Int'l Conference in Evolutionary Multi-Criteria Optimizaton*, 2003, pp. 8–11.

[4] N. Hayashida and H. Takagi, "Acceleration of ec convergence with landscape visualization and human intervention," *Soft Computing*, vol. 1, pp. 245–256, 2002.

[5] A. Ultsch, G. Guimaraes, D. Korus, and H. Li, "Knowledge extraction from artificial neural networks and applications," in *TAT/WTC93*, 1993, pp. 194–203.

[6] M. M. V. Hulle, *Self Organization Map – theory, design and application – (transulated by Tokutaka and Fujimura)*. Tokyo, Japan: Kaibundo Press, 1999.

[7] H. Tokutaka, S. Kishida, and K. Fujimura, *Application of Self Organizing Map – Visualization of multidimensional data to two dimensional space –*. Tokyo, Japan: Kaibundo Press, 1999.

[8] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multi-objective genetic algorithm: Nsga-ii," *IEEE Trans. on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[9] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley and Sons, LTD, 2001.

[10] K. Urushibata and M. Oyabu, "Development of the spherical som and its property," in *17th Fuzzy System Symposium collected papers*, Japan, 2001, pp. 339–342.

[11] D. Nakatsuka and M. Oyabu, "Usefulness of spherical som for clustering," in *19th Fuzzy System Symposium collected papers*, Japan, 2003, pp. 67–70.

[12] H. Kurosawa, K. Fujimura, M. Ohkita, H. Tokutaka, and M. Ohyabu, "The usefulness of spherical self-organizing maps," in *19th Fuzzy System Symposium collected papers*, Japan, 2003, pp. 71–74.

[13] H. Ishibuchi and Y. Shibata, "Mating scheme for controlling the diversity-convergence balance for multiobjective optimization," in *2004 Genetic and Evolutionary Computation Conference Part 1*, Japan, 2004, pp. 1259–1271.

[14] N. Ueno and T. Furuhashi, "A nurse scheduling support system with nurse-in-chief's know-how," in *2004 14th Intelligent System Symposium*, Japan, 2004, pp. 174–177.