A Visualization of Genetic Algorithm Using the Pseudo-color

Shin-ichi Ito[1], Yasue Mitsukura[1], Hiroko Nakamura Miyamura[1],
Takafumi Saito[1], and Minoru Fukumi[2]

[1] Graduate School of Bio-Applications & Systems Engineering,
Tokyo University of Agriculture and Technology
2-24-16, Naka, Koganei, Tokyo, 184-8588, Japan
{ito_s,mitsu_e,miyamura,txsaito}@cc.tuat.ac.jp
[2] The University of Tokushima
2-1, Minami-Josanjima, Tokushima, 770-8506, Japan
fukumi@is.tokushima-u.ac.jp

Abstract. In this paper, we propose a visualization method to grasp the search process and results in the binary-coded genetic algorithm. The representation, the choices of operations, and the associated parameters can each make a major difference to the speed and the quality of the final result. These parameters are decided interactively and very difficult to disentangle their effects. Therefore, we focus on the chromosome structure, the fitness function, the objective function, the termination conditions, and the association among these parameters. We can indicate the most important or optimum parameters in visually. The proposed method is indicated all individuals of the current generation using the pseudo-color. The pixels related a gene of the chromosome are painted the red color when the gene of the chromosome represents ‘1’, and the pixels related to one are painted the blue color when one represents ‘0’. Then the brightness of the chromosome changes by the fitness value, and the hue of the chromosome changes by the objective value. In order to show the effectiveness of the proposed method, we apply the proposed method to the zero-one knapsack problems.

Keywords: binary-coded genetic algorithm, zero-one knapsack problem, visualization, pseudo-color.

1 Introduction

It can be easy to quickly set up a genetic algorithm (GA), which is a search paradigm that applies ideas from evolutionary genetic operations (natural selection, crossover, mutation) in order to search an optimum solution or a quasi-optimal solution, but analyzing the results to discover whether or not the process is efficient or could be improved is often extremely difficult. The representation, the choices of operations, and the associated parameters can each make a major difference to the speed and the quality of the final result. These parameters are decided interactively and very difficult to disentangle their effects [1], because

M. Ishikawa et al. (Eds.): ICONIP 2007, Part II, LNCS 4985, pp. 444–452, 2008.

the user sets a chromosome structure of an individual, a fitness function, an objective function, genetic operation parameters, and the termination conditions.

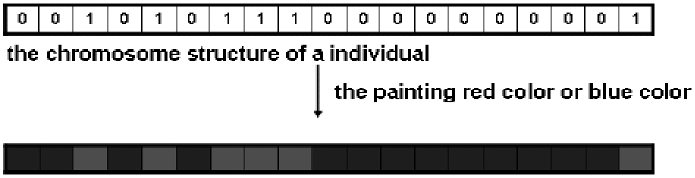
Generally, the objective function and the fitness function are the same function, but there is the case that the objective function and the fitness function are the difference, because there is the case that the fitness function includes the user's knowledge gave as the penalty and has the multi-objective function. Furthermore, keeping a full record of everything that occurs during a GA run produces large quantities of data that can not be analyzed conveniently by hand. Ideally, the full recording data displays on only one-frame for carrying out the discussion of setting parameters and evaluating the fitness function and the penalty. We focus on the chromosome structure, the fitness function, the objective function, association among these parameters, and whether termination conditions have been satisfied. Then we indicate these on only one-frame. However, it is very difficult that the some parameters are indicated simultaneously. We proposed the visualization method to grasp the search process, the search result, and setting parameter associations. Then, the process of the genetic search of the GA can be visualized. We can indicate the most important or optimum parameters in visually. Therefore, we indicate all individuals' chromosomes of the current generation by using the pseudo-color; The pixels related a gene of the chromosome are painted the red color when the gene of the chromosome is '1', and the pixels related to one are painted the blue color when one is '0', the brightness of the chromosome changes by the fitness value, and the hue of the chromosome changes by the objective value. In order to show the effectiveness of the proposed method, we apply the proposed method to the zero-one knapsack problems (KP).

This paper is organized in the following manner. First, in the section 2, we introduce the zero-one knapsack problem. In the section 3, we explain the visualization method of GA based on the pseudo-color. In the section 4, we describe computer simulations applying the KP for showing the effectiveness of the proposed method. In the section 5, we discuss the proposed method and the computer simulation results. Finally, we indicate the conclusions and the feature works.

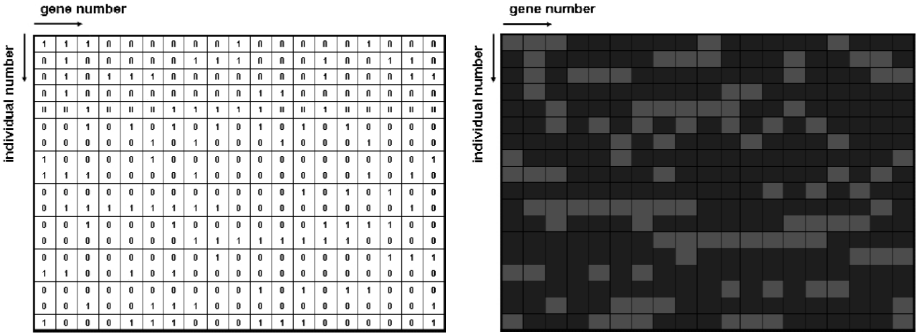
2 Knapsack Problem

The zero-one knapsack problems seek to place objects in a knapsack so as to maximum the total value of the objects without overfilling the knapsack. We are given n objects, each with a value v_i and weight w_i, and a knapsack with capacity C, and we seek a selection of objects for knapsack with maximum total value but with total weight no greater than C. That is, if n binary variables x_i indicate the inclusion ($x_i = 1$) or exclusion ($x_i = 0$) on each object. Formally, the KP can be stated as follows:

$$\begin{aligned} &maximize \sum_{i=1}^{n} x_i v_i, \qquad (1) \\ &subject\ to \sum_{i=1}^{n} x_i w_i \leq C, \quad v_i \geq 0, \quad w_i \geq 0, \quad i = 1, ..., n. \end{aligned}$$



(a) The sample result of a chromosome painted red color and blue color.



1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
0	1	0	0	0	0	0	1	1	1	0	0	0	1	0	0	1	1	0
0	1	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	1
0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
0	0	1	0	0	0	1	1	1	1	1	0	0	1	0	0	0	0	0
0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	0	0
0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	1	0	0	0
1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0
0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	0
0	0	1	0	0	0	1	0	0	0	0	0	0	1	1	1	1	0	0
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	1
1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0
0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	0	1

(b) All chromosomes of a current generation.

(c) The sample result of all chromosomes painted red color and blue color.

Fig. 1. The chromosome structure and the sample result of the individual chromosomes of a current generation painted red color and blue color

3 Methods

3.1 Genetic Encoding

A simple genetic encoding scheme for the KP is as follows. Let each bit represent the inclusion or exclusion of one of the n objects from the knapsack. Note that it is possible to represent infeasible solutions by setting so many bits to '1' that the weight of the corresponding set of objects overflows the capacity of the knapsack.

3.2 Objective Function and Fitness Functions

The objective function is objective assessment (e.g., the recognition accuracy in the pattern classification problem, the total value in the KP). The fitness function is an evaluation function for searching the optimum solution or the quasi-optimum solution. Generally, the objective function and the fitness function are the same function. However there is the case that the objective function and the fitness function are the difference because there is the case that the fitness function includes the user's experimental knowledge and has some objective functions.

The objective function of the KP is shown Eq. (1). The case that the objective function (O) and the fitness function are the same function, the fitness function (F) is as follows:

$$maximize\ F = O = \sum_{i=1}^{n} x_i v_i, \tag{2}$$

$$subject\ to\ W = \sum_{i=1}^{n} x_i w_i \leq C, \quad v_i \geq 0, \ \ w_i \geq 0, \ \ i = 1, ..., n.$$

The case that the objective function and the fitness function are the difference, the fitness function has a penalty function as follows:

$$maximize\ F = \alpha O - \beta P, \tag{3}$$

$$subject\ to\ W = \sum_{i=1}^{n} x_i w_i \leq C, \quad v_i \geq 0, \ \ w_i \geq 0, \ \ i = 1, ..., n,$$

where α means the weight of the objective function, β means the weight of the penalty function (P). It is difficult that the user determines how to give the penalty and the weight value. Therefore the penalty is determined the user's knowledge and the weight value is determined experimentally. Then it is the difficult to evaluate the penalty, the weight value, and whether to search the optimum solution and/or the quasi-optimal solution.

3.3 Visualization of Genetic Algorithm

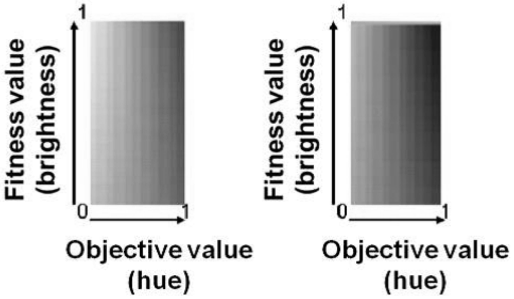
Keeping a full record of everything that occurs during a GA run produces large quantities of data that can not be analyzed conveniently by hand. Ideally, the full recording data displays on only one frame for carrying out the discussion of setting parameters of the GA and evaluating the fitness function and the penalty. We focus on the chromosome structure, the fitness function, the objective function, the association among these parameters, and whether to be satisfied the termination conditions, and then we indicate these on one-image. However, it is very difficult that the some parameters are indicated simultaneously. Therefore, we indicate all individuals' chromosomes of the current generation by using the pseudo-color. Then any generations indicate on only one-frame.

The method for visualization in the GA is as follows:

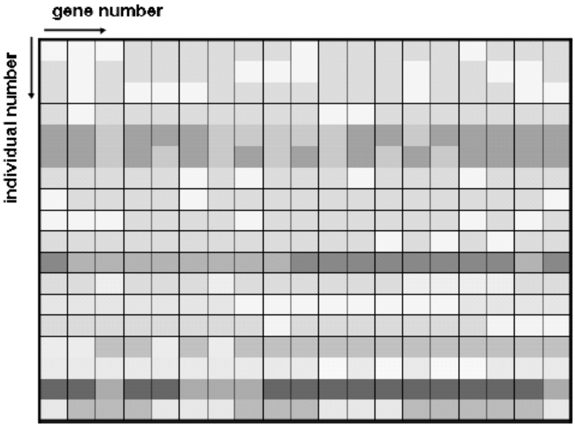
a) The image of $n * m$ is used when the length of chromosome is n and the number of the individual is m. A pixel of the image is related to a gene of the chromosome. The pixels related a gene of the chromosome are painted the red color when the gene of the chromosome is '1', and the pixels related to one are painted the blue color when one is '0' shown in Fig.1.

b) Fig.2 shows the sample result of the visualization using the pseudo-color. The brightness (range: $0 \leq Red_{brightness} \leq 40$, $120 \leq Blue_{brightness} \leq 160$) of the chromosome changes by the fitness value. Moreover the hue (range: $0 \leq Hue \leq 128$) of the chromosome changes by the objective value. In this paper, the range of the objective value and fitness value is 0.0 to 1.0 for having related the objective value and the fitness value to the brightness change and the hue change. We make the visualization image changed hue and brightness shown in Fig.2(b).

c) Any generations are indicated on one-frame shown in Fig.3.



(a) The pseudo-color



(b) All chromosomes of a current generation changed brightness and hue based on the objective value and fitness value

Fig. 2. The sample result of the visualization using the pseudo-color

4 Computer Simulations

In order to show the effectiveness of the proposed method, we solve the KP using three fitness functions (Eqs.(4)-(6)). The range of the fitness functions is from 0.0 to 1.0.

$$O = F_1 = (1.0 - \frac{T}{1.0+\sum_{i=1}^{N} x_i v_i})^3 \tag{4}$$

$$subject\ to\ \ W = \sum_{i=1}^{n} x_i w_i \leq C, \quad v_i \geq 0, \ \ w_i \geq 0, \ \ i = 1, ..., n$$

$$F_2 = (1.0 - \frac{T}{1.0+\sum_{i=1}^{N} x_i v_i} - \frac{NumGene`1'}{ChromoLen})^3 \tag{5}$$

$$subject\ to\ \ W = \sum_{i=1}^{n} x_i w_i \leq C, \quad v_i \geq 0, \ \ w_i \geq 0, \ \ i = 1, ..., n$$

$$F_3 = (0.9 * (1.0 - \frac{T}{1.0+\sum_{i=1}^{N} x_i v_i}) - 0.1 * (\frac{NumGene`1'}{ChromoLen}))^3 \tag{6}$$

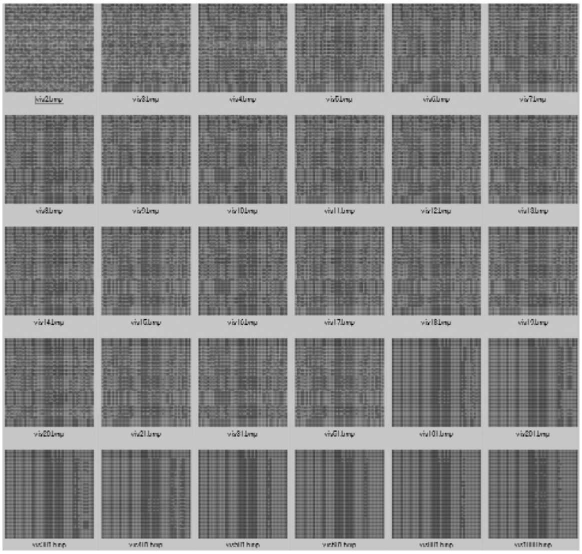


Fig. 3. The sample of the visualization frame indicated any generations

Table 1. The value and weight of the objects

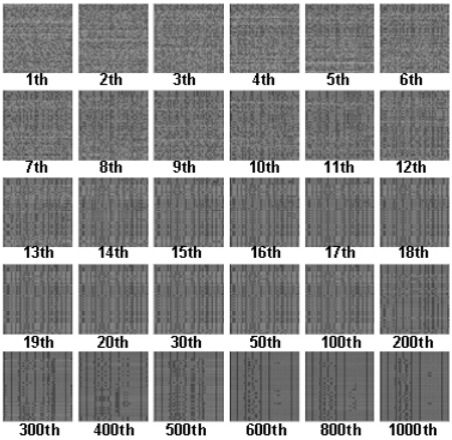
object number	value	weight	object number	value	weight	object number	value	weight
(1)	1	24	(18)	9	6	(35)	15	24
(2)	14	21	(19)	18	0	(36)	13	28
(3)	4	29	(20)	4	13	(37)	15	1
(4)	13	19	(21)	14	26	(38)	16	19
(5)	10	2	(22)	3	11	(39)	5	23
(6)	6	1	(23)	7	5	(40)	3	27
(7)	13	9	(24)	12	26	(41)	17	25
(8)	16	17	(25)	5	29	(42)	19	2
(9)	19	11	(26)	15	22	(43)	0	16
(10)	19	13	(27)	19	4	(44)	13	3
(11)	19	26	(30)	11	27	(45)	2	13
(12)	3	10	(31)	14	27	(46)	6	23
(13)	0	1	(28)	14	4	(47)	12	25
(14)	9	11	(29)	10	23	(48)	0	19
(15)	4	5	(32)	4	3	(49)	9	2
(16)	8	18	(33)	6	5	(50)	12	6
(17)	0	0	(34)	10	25			

$$subject\ to\ \ W = \sum_{i=1}^{n} x_i w_i \le C, \quad v_i \ge 0, \ \ w_i \ge 0, \ \ i = 1, ..., n,$$

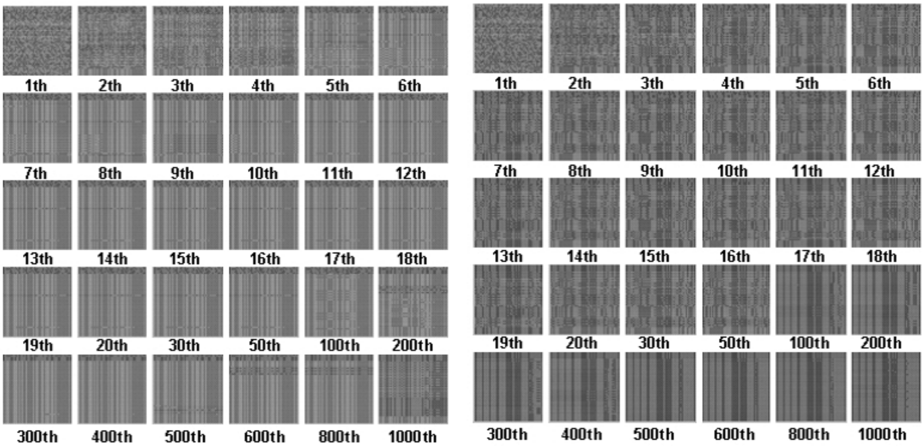
where O means the objective function, F_n means the fitness functions, $NumGene$'$1'$ indicates the number of the gene that is '1' of an individual chromosome, $ChromoLen$ indicates the length of a chromosome.

Table 2. The parameters of the GA

The number of generations	1,000
The number of individuals	50
The length of chromosome	50
The rate of elite	0.1
The two-point crossover rate	0.8
The mutation rate	0.1



(a) The sample result of the case that the objective function and the fitness function are the same function



(b) The sample result of the case that the fitness function included the penalty not satisfied the user's desire

(c) The sample result of the case that the fitness function included the penalty satisfied the user's desire

Fig. 4. The sample of the simulation results

The genetic operations of the GA are the two-point crossover, the mutation, and the elite strategy. Table 1 shows the value and weight of the object of the KP, and Table 2 shows the parameters of the GA. The sample results of the computer simulation shown in Fig.4; (a) the case that the objective function and the fitness function (Eq.(4)) are the same function, (b) the case that the objective function and the fitness function are the difference and the fitness function (Eq.(5)) does not be given the penalty appropriate for being satisfied user's desire, (c) the case that the fitness function (Eq.(6)) is given the penalty appropriate for being satisfied user's desire. The generation number indicated the frame shown in Fig.4 is from 1 to 20, 30, 50, 100, 200, 300, 400, 500, 600, 800, and 1000.

5 Discussions

Based on the results shown in Fig.4, it is reasonable to support that it is understood intuitively that the chromosome structure, the fitness function, the objective function, association among these parameters on each generations. On the other hands, we think the validity of the fitness function is confirmed by seeing the colors of all generations and each generation, because if the colors are cloudy and light then the user knows the fitness function is not set appropriately shown in Fig.4(b) and if the colors are clear and dark then the user thinks the fitness function is set appropriately shown in Fig.4(c). Then we think the user is able to confirm whether to be satisfied the termination conditions set by the user, because if the termination conditions are satisfied then the color-gradation of all chromosomes on any generation is similar or same from top (individual number is 1) to bottom (individual number is 1000) shown in Fig.4(c). Moreover, if the color-gradation from top to bottom is not similar shown in Fig.4(b) then the user thinks that the termination conditions may be not able to be satisfied, because the color-gradation of the upper part and the low part is not similar for giving strongly the penalty of Eq.(5).

6 Conclusions and Future Works

We propose the method of the visualization in the search process and the search result of a binary-coded genetic algorithm. We focus on the chromosome structure, the fitness function, the objective function, association among these parameters, and whether to be satisfied the termination conditions. The proposed method makes the visualization image changed hue and brightness for showing the relationship among the chromosome structure, the fitness function, and the objective function. The chromosome of the individual having the most vivid color and the strongest contrast red color and blue color shows the highest fitness value and the optimum solution or the quasi-optimum solution. Furthermore, any generations are indicated on one-frame for showing the generation change that has the fitness value change, the objective value change, the searching condition of the optimum solution or the quasi-optimum solution, and whether to be satisfied the termination conditions. Finally, the performance of the proposed method was evaluated applying to the zero-one knapsack problem. It is understood intuitively that the proposed method shows the chromosome structure, the fitness function, the objective function, the association among these parameters, and whether to be satisfied the termination conditions.

We will involve efforts to improve the visualization techniques, except of using the pseudo-color, to further develop a real-coded genetic algorithm and an interaction genetic algorithm, to be used many users for evaluating the proposed method in the future works.

References

1. Hart, E., Ross, P.: Gavel-A New Tool for Genetic Algorithm Visualization. IEEE Transaction on Evolutionary Computation 5(4), 335–348 (2001)
2. Eick, S.G., Steffen, J.L., Summer, E.E.: Seesoft-A Tool for Visualization Line Oriented Software Statistics. IEEE Transaction on Software Engineering 18, 957–968 (1992)
3. Simoes, A., Costa, E.: An Evolutionary Approach to the Sero/One Knapsack Problem: Testing Ideas from Biology. In: Kurkova, V., Steele, N., Neruda, R., Karny, M. (eds.) Proceedings of the Fifth International Conference on Neural Networks and Genetic Algorithms (ICANNGA 2001), Prague, Czech Republic, April 22-25, pp. 236–239. Springer, Heidelberg (2001)
4. Jones, T.: Crossover, macromutation, and population-based search. In: Eshelman, L. (ed.) Proceedings of the 6th International Conference on Genetic Algorithms, pp. 73–80. Morgan Kaufmann, San Mateo (1995)
5. Shine, W., Eick, C.: Visualization the evolution of genetic algorithm search processes. In: Proceedings of 1997 IEEE International Conference on Evolutionary Computation, pp. 367–372. IEEE Press, Piscataway (1997)
6. Olsen, A.L.: Penalty Function and the Knapsack Problem. In: Fogel, D.B. (ed.) Proceedings of the 1st International Conference on Evolutionary Computation 1994, Orlando, FL, pp. 559–564 (1994)
7. Gordon, V., Bohm, A., Whitley, D.: A Note on the Performance of Genetic Algorithms on Zero-One Knapsack Problems. In: Proceedings of the 9th Symposium on Applied Computing (SAC 1994), Genetic Algorithms and Combinatorial Optimization, Phoenix, Az (1994)