A Constructive Approach to Image Understanding

Kei OHNISHI and Kaori YOSHIDA

Kyushu Institute of Technology

680–4 Kawazu, Iizuka-shi, Fukuoka, 820-8502, JAPAN

k_ohnisi@pluto.ai.kyutech.ac.jp kaori@ai.kyutech.ac.jp

***Abstract*— This paper proposes a constructive approach to understanding a given static structure, which tries to understand the given static structure through interpreting a process of constructing it. Two things are necessary for crystallizing this concept. One is a method for generating structures which provides an interpretable process of generating a structure. Another is an optimization method which adjusts parameters values of a method for generating structures for that method to achieve construction of a given structure. A constructive method realized in this paper combines a method for generating structures inspired by biological development with a real-coded genetic algorithm as an optimization method. The realized method is applied to image understanding as a sort of understanding static structures, and the results show the prospects of the proposed concept.**

I. Introduction

Human beings have been obtaining the descriptions of nature based on the observation and the analysis of nature, which is called natural science. The descriptions obtained have been used for the control of nature as well as nature understanding. Under the safety guaranteed by those efforts, research fields that try to make society better have emerged. Engineering is one of the representatives.

Many research fields emerged in the safe society that the natural science promises have inherited the natural scientific approach, which understands objects based on their observation and analysis. Meanwhile, research fields of artificial life [6][10] and complex system [15][7][4][5] came out in the late 1990's. An approach to understanding objects in those research fields is different from the natural scientific approach. In those research fields, people try to understand nature not analytically but synthetically. That is, they first try to construct observed natural behaviors using some models somehow, and then understand nature through interpretations of obtained processes and rules for constructing the observed natural behaviors. This is generally called a *constructive approach*.

The constructive approach constructs a natural object's behavior by means of local interactions among components of the whole object. In this case, it could be a problem that the obtained process of the construction is not the sole process for the observed behavior and interpretations of the obtained interaction rules are ambiguous. However, if the observed natural object's behavior can be constructed somehow, a set of the obtained rules for the construction is at least able to be one of the mechanisms behind the observed behavior even if they are not real mechanisms, so that they should help us understand the observed behavior.

As for human society, it is getting more and more informational and the Internet is getting worldwide, which is so-called information society. Now human-human, human-machine, and machine-machine communications beyond physical distances are possible. In a vast virtual space built on the Internet, there exists a variety of electronic data in which people get interested as in the real world. We can also use the two approaches mentioned above, which are the natural scientific approach and the constructive approach, to understanding electronic data.

For instance, most of the technologies frequently used in information society, such as information retrieval techniques, rely on analytical descriptions of digital contents as same as the natural scientific approach. Moreover, since analytical descriptions of digital contents are easy for computers to handle but difficult for humans to interpret intuitively, multimedia contents retrieval systems using natural language as queries have been proposed [16][14]. In those systems, mapping between natural language and analytical descriptions of multimedia contents enables us to retrieve multimedia contents in such a way.

Studies on electronic data processing and generating with the constructive approach have also increased. Artificial life approach has been applied to image processing [3]. Interactive evolutionary computation (IEC) has been applied to generating multimedia contents [13]. IEC is an optimization technique which involves human as an evaluation function for a given optimization problem, so that IEC can optimize system parameters along human evaluation criteria without analyzing and obtaining the human evaluation criteria.

The objectives of this paper are to propose a concept of the constructive approach to understanding static structures including electronic data such as image, movie, music, and text, and to realize a constructive method for understanding static structures along the proposed concept. A static structure means a structure that does not vary whenever it is referred to. This approach firstly tries to construct a given static structure using some model somehow and then we utilize interpretations of the obtained process of the construction for understanding the given structure.

This paper is organized as follows. In section II, we proposes the concept of the constructive approach to understanding static structures. Section III explains the constructive method realized along the proposed concept. In section IV, the realized method is applied to image understanding as a sort of understanding static structures. Finally, we summarize our results and draw our conclusions in section V.

II. Concept

The constructive approach to understanding static structures is that an interpretable process of construction for a give structure is firstly obtained and then we try to understand the given structure through interpretations of the obtained process of constructing the given structure. On the other hand, in case that some approach can perfectly construct a given structure but a process of constructing the given structure is hard for humans to interpret, the approach just tells us the fact that it can model the given structure. Furthermore, the difference between conventional data modeling and mining techniques and the proposed approach is that while conventional data modeling and mining techniques generate a static structure different from an original structure to understand the original structure, the proposed approach uses a process of constructing an original structure to understand the original structure.

This paper assumes static structures generated by a series of procedures. When we generate a whole structure in a finite space by a series of procedures, we need to consider three things below.

(1) A rough picture of a whole structure,
(2) Methods for generating portions of the whole structure,
(3) An order of applying the methods.

For example, when we paint a picture based on our imagination or what we see, in both cases, we need to (1) get a rough picture, (2) decompose the rough picture into its portions, and (3) apply methods for painting the portions one after another. In this case, since portions painted earlier become spatial restrictions against portions painted later, this way is more complicated than a way that paints portions of a whole picture independently.

When there is a method for generating structures that considers the three procedures mentioned above and the method can generate a variety of structures by changing its parameters, there is a possibility that the method can construct a given static structure by adjusting its parameters values. Furthermore, if we can construct a given structure with such a method not perfectly but somewhat and a process of constructing the given structure is easy for us to interpret, the process helps us understand the given structure deeper even if the process is merely one of the possible processes.

III. A Constructive Method

To realize the concept proposed in the last section, we need to prepare two methods below. We call a method which combines these two methods a *constructive method*.

(a) A method for generating structures which provides an interpretable process of generating a structure.
(b) An optimization technique which optimize parameters values of a method for generating structures.

We explain those two methods below.

A. A Method for Generating Structures

We use a method inspired by biological development [1][11] as a method for generating static structures which provides an interpretable process of generating a static structure.

Biological development is a process in which a mother cell grows up to be an adult body while repeating cell division and specialization. A genome in the mother cell is a design information for development. Since the genome in the mother cell is accurately copied into a new cell one after another, biological development is based only on the genome in the mother cell.

A mechanism which plays an important role in an early stage of development is proteins hierarchically diffusing among cells. Each cell gradually acquires information on what organ it will finally compose according to density of the diffusing proteins. The information that each cell acquires is called *positional information*. For instance, proteins diffusing among cells earlier give the cells rough information on their final organs, and proteins diffusing later give them detailed information on those. In this way, proteins hierarchically diffusing among cells place the cells on correct positions while producing various cells. This phenomenon is called *pattern formation*.

The method for generating structures used in this paper is a feedback system whose main strategy to generate structures is based on the principal of the pattern formation. The analogies between the mechanism of biological development and the method used in this paper are listed in the below (a)'s and (b)'s, respectively, and they are illustrated in Fig. 1. The numbers attached in the figure correspond to the following seven analogies.

1. Body and structure: (a) An adult body consists of various cells. (b) A structure consists of various components.

2. Design information: (a) Only a genome in a mother cell is design information in development. (b) A structure is generated according only to an array storing parameters values, which is called a *character code* thereafter.

3. Hierarchy: (a) A body is hierarchically formed from rough to detailed. The procedure in one hierarchy follows 4(a)→5(a)→6(a). (b) Structure is hierarchically generated from rough to detailed. The method used in this paper is realized as a feedback system by corresponding one hierarchy to its main processing part. The main processing part follows 4(b)→5(b)→6(b) and the feedback processing part is 7(b).

4. Positional information: (a) The protein diffusing among all or some cells gives them positional information. (b) There is a mechanism to generate positional information among all or some components.

5. Response to positional information: (a) Each cell acts in response to its own positional information, for example, generating special proteins. (b) Each components acts in response to its own positional information.

6. Memorizing positional information: (a) The state of each cell after acting in response to its positional information is memorized to reach its final state. (b) The state of each component after acting in response to its positional information is memorized to reach its final state.

7. Feedback: (a) Proteins that provide current positional information are produced based on proteins that provided past positional information. Each cell is gradually led to a final organ by a history of given various positional information. (b) A mechanism that generates current positional information using past positional information is prepared. An option of each component's final state gradually becomes narrower according to a history of given various positional information.

We will explain the method used in this paper using a simple example here. The task of the method is to generate real numbers ($y \in \boldsymbol{R}$) on one dimensional integer coordinates in a finite range ($X = \{x \in \boldsymbol{Z} \mid x \in [x_1, x_n]\}$). In this case, distribution of real numbers on the integer coordinates is a structure that the method generates. Since the method is a feedback system, it consists of main processing and feedback processing parts. Parameters values that each processing part uses are stored int a character code in their used order.

Main processing part

The main processing part is a series connection of the following three sub-processing parts ((1)→(2)→(3)).

(1) Generating positional information

This sub-processing part generates distribution of real numbers on the integer coordinates as proteins diffusing among cells in biological development. A real number on each integer coordinate can be regarded as positional information. In this example, the distribution of real numbers is generated by the normalized sum of functions with Gaussian shape (see Fig. 2(1)). Co-

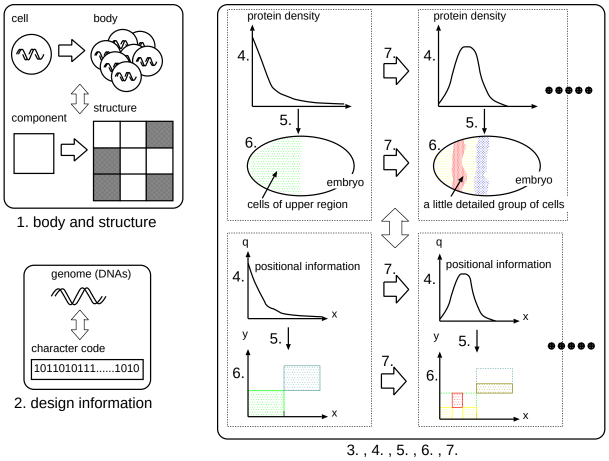


Fig. 1. Analogies between biological development and the method for generating static structures used in this paper.

efficient, expectation, and standard deviation of each function with Gaussian shape are parameters of the method, and their values are stored in the character code.

(2) Responding positional information

This sub-processing part decides how each component (integer coordinate) acts in response to its positional information given by the sub-processing part (1). Concretely, it classifies the integer coordinates into several groups of coordinates (see Fig. 2(2)). Each group is assigned a label. The maximum number of kinds of labels has to be provided prior to starting to generate a structure. Which labels are assigned to classified groups of coordinates are parameters of the method.

(3) Determining states of components

This sub-processing part converts the label assigned to each coordinate in the sub-processing part (2) into a range of y values that each coordinate can finally have (see Fig. 2(3)). This processing corresponds to the fact that a history of positional information that each cell has taken relates to determination of a final organ that it will compose. Rules for converting a label into a range of y values are parameters of the method.

Feedback processing part

The sub-processing parts (1) and (3) include feedback processing.

The method for generating structures used in this example first defines a set of components which are integer coordinates, and then starts generating a structure by generating distribution of real numbers on the components using functions with Gaussian shape. To start the second main processing, new distribution of real numbers on the components has to be generated as in the first main processing. To generate the distribution of real numbers in the sub-processing (1) from second time downward, expectations of functions with Gaussian shape are firstly obtained as intersections between the last normalized sum of functions with Gaussian shape and lines paralleling the axis in which components exist, and then a new normalized sum of functions with Gaussian shape is generated using the obtained expectations and the other parameters values from the character code (see Fig. 3).

The output from the sub-processing part (3) is a range of y values that each component can finally take. The sub-processing part (3) outputs not a range of y values independent of a previous ranges but a range down which the last range is narrowed. The last ranges of y values that all the components have are feedback data to the sub-processing part (3). Therefore, output from the sub-processing part (3) always stands for the outline of a final structure, and the change of outputs from it represents a process of how a structure was getting complicated.

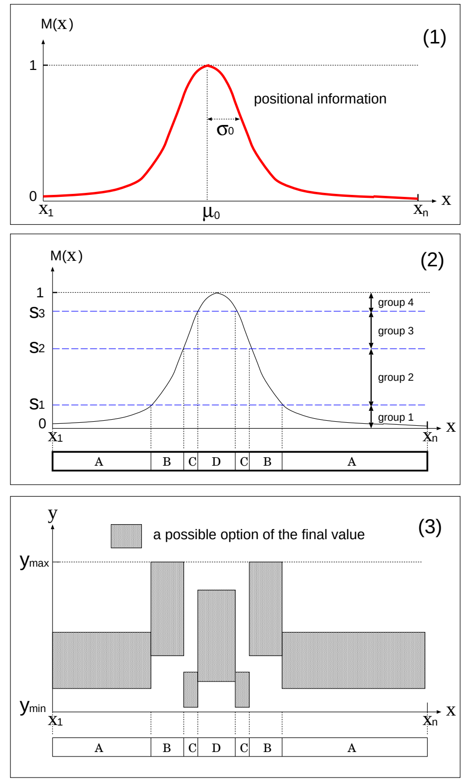


Fig. 2. An example of a main processing.

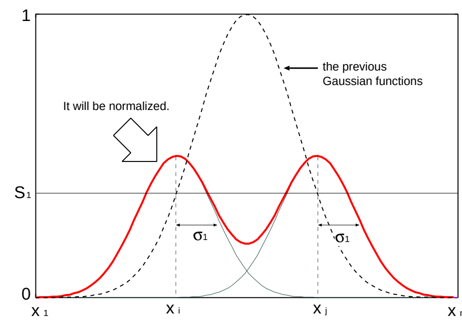


Fig. 3. Data feedback to the sub-processing part (1) and generation of new positional information.

The change of outputs from the sub-processing part (3) including the feedback processing is the most important for human to interpret a process of generating a structure because it is a process of generating a structure itself. In the example of generating a structure above, there was no interactions among components. However, even if interactions among components are used, as long as a state of each component at some moment is affected by its past state, following the change of outputs from the sub-processing part (3) should be meaningful for interpreting a process of generating a structure which consists of such components

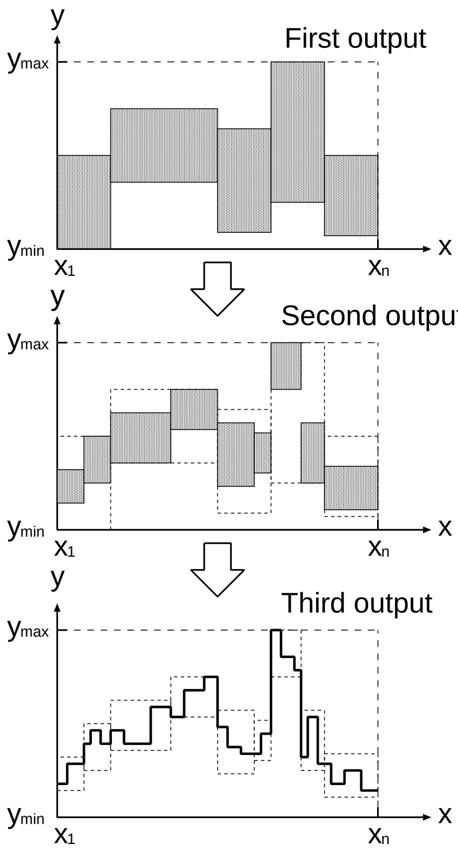


Fig. 4. An example of transition of outputs from the sub-processing part (3).

B. An Optimization Method

A role of an optimization method here is to adjust the parameters values of the method for generating structures described in the last section in order for the method for generating structures to achieve construction of a given structure with high accuracy. The parameters are represented as a one dimensional array, which is the character code, and by giving a set of concrete values to the array, one structure is generated.

We have to define the criterion of good construction of a give structure based on the differences between a given structure and a constructed structure, and quantify the criterion to be available for an optimization method, as well. We will call values representing the quantified criterion *fitness values* thereafter. Under the defined criterion, we can use any optimization methods suitable for a given type of object variables. If we do not have any knowledge about distribution of fitness values on an object variables space, black-box optimization algorithms, which do not demand a prior knowledge about a given optimization problem, should be appropriate. Genetic algorithm (GA) [8][9] is one of the representative black-box optimization algorithms.

What we should pay attention to when we optimize the parameters values of the method for generating structures explained in the last section is that as shown in Fig. 4 the sub-processing part (3) outputs an option of final values that each coordinate can finally take and the option is gradually narrowed down with iterations of the main processing. For instance, if output from the sub-processing part (3) does not include some part of a given structure at some moment, we loose a possibility of constructing a given structure perfectly at this moment. Therefore, we should take into account not only the difference between a given structure and a final structure from the sub-processing part (3) but also the difference between a given structure and a medium structure from it.

IV. Image Understanding

In this section, we try to understand a simple image by means of the constructive method for understanding static structures described in section III.

A. Task

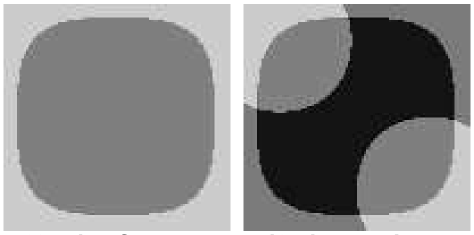
In order for the constructive method to construct an objective image that we want to understand, an optimization method adjusts the parameters values of the constructive method. As a result, we obtain a process of constructing the objective image.

An image that the constructive method will construct is a gray-scaled image with 100×100 pixels. The constructive method looks on the objective image that we want to understand as distribution of integer values from 0 to 255 on two dimensional integer coordinates within a 100×100 square area. The objective image is generated by the same method as for constructing the objective image, so that the constructive method can strictly construct the objective image if an optimization method successfully adjusts the parameters values of the constructive method. The objective image and its construction process is shown in Fig. 5. Since medium structures obtained from the sub-processing part (3) represent not values but ranges in which final values are determined on coordinates and they can not be drawn unless we select one value from each range, the medium structure shown in Fig. 5(a) was drawn with representative values within the ranges. The representative values were algorithmically determined and not always the middle points of the ranges, but roughly speaking, they were around the middle points in most cases.

As shown in Fig. 5(b), nothing that we can recognize as conception beyond geometric figures is drawn in the objective image, so it is hard to see the process of generating the objective image from the viewpoint of how objects recognized as conception beyond geometric figures are generated. Here we first compare the process of generating the objective image (Fig. 5(a)(b)) with the process of constructing the objective image by the constructive method, though the medium structure shown in Fig. 5(a) is not used for calculating a fitness value to each construction. Then if those two processes are similar to each other, we think that the process obtained by the constructive method helps us understand the objective image. In fact, if obtained processes of constructing the objective image by the constructive method are helpful for understanding the objective image, we could say that the constructive method is useful no matter what fitness values of constructed images. However, in this paper, we do not say the usefulness of the proposed concept in such a way.

B. Configurations

We use an extended method described in section III-A to construct the objective image, which is able to generate distribution of values on two dimensional integer coordinates. The extended method uses Gaussian shaped functions with two variables to generate distribution of values on the two dimensional integer



(a) The first output from the sub-processing part (3) (medium structure).

(b) The second output from the sub-processing part (3) (final structure)

Fig. 5. The process of constructing the objective image that we try to understand.

coordinates in the sub-processing part (1), which is the positional information. The sub-processing part (1) sums up and normalizes four functions with Gaussian shape. Moreover, the number of planes paralleling the two dimensional space, which correspond to S_i in Fig. 2(2) and Fig. 3, is one. The sub-processing part (2) gives a group of integer coordinates one of the two kinds of labels. Output from the sub-processing part (3) is distribution of real numbers within 0 to 1 or a finite range within 0 to 1 on the integer coordinates. The real number and the finite real range are converted into integer within 0 to 255 and a finite integer range, respectively. The number of iterations of the main processing is two, and therefore the number of the feedback processing is one.

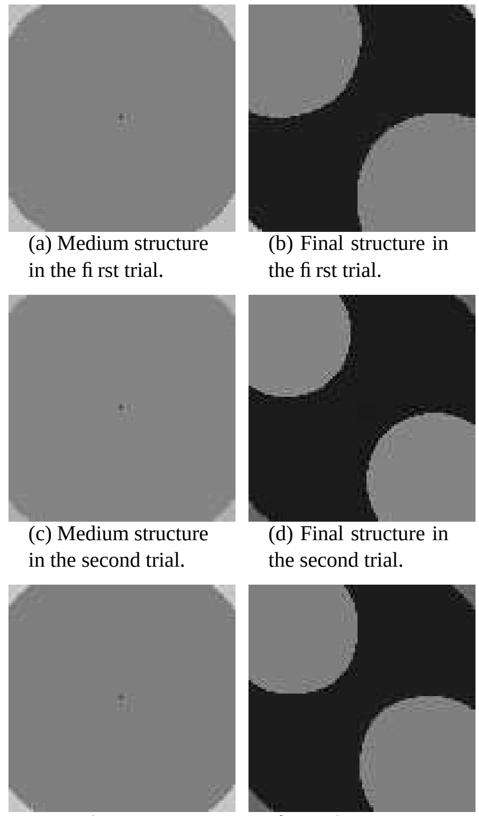
The parameters that an optimization method optimizes are real numbers within 0 to 1, and the number of them is 19. The actual number of the parameter is more than that. However, since we saw through preliminary experiment that it was difficult for an optimization method to optimize the actual parameters, we decreased the actual parameters in number by fixing parameters values related to the sub-processing part (1) except parameters values corresponding to S_1 in Fig. 3. The parameters values are converted into values in an appropriate range or integers when they are used.

We use a real-coded GA as an optimization method which optimizes the parameters values of the method for constructing the objective image. The used real-coded GA consists of minimal generation gap model (MGG) [12] and BLX-α crossover operator [2]. It does not use mutation operator. Its population size is 10000 and its stop condition is 5×10^5 fitness evaluations. A GA individual is a set of the parameters values of the method for constructing the objective image.

The fitness value of a GA individual is calculated by comparing both of the first and the second outputs from the processing part (3) with the objective image shown in Fig. 5(b). A portion of the entire fitness value obtained by comparing the first output with the objective image is in proportion to how many the first output includes the objective image, and a portion of the entire fitness value obtained using the second output is in proportion to the number of times that each pixel in the second output has the same integer as the corresponding pixel in the objective image has. The ratio of the first output's and the second output's maximum contribution to the entire fitness value is 4 to 1.

C. Results and Discussions

We applied the constructive method to the objective image three times. The constructed image with the highest fitness value in each trial is shown in Fig. 6 together with its constructing process. The first outputs from the sub-processing part (3) could not be drawn because they were a set of ranges in which final values would be determined, so that we drew the first outputs with the representative values, which were mentioned in section IV-A.



(a) Medium structure in the first trial.

(b) Final structure in the first trial.

(c) Medium structure in the second trial.

(d) Final structure in the second trial.

(e) Medium structure in the third trial.

(f) Final structure in the third trial.

Fig. 6. The processes of constructing the objective images by the realized constructive method.

Comparing Fig 5 and Fig. 6, we can see that the constructive method constructed the objective image not perfectly but roughly. In addition, although we did not use the medium structure in the process of generating the objective image for calculating fitness values of the GA individuals, the processes of constructing the objective image were also similar to the process of generating the objective image. Therefore, under the idea that accurate construction of the objective image helps us understand the objective image, the obtained results show the possibility of the proposed concept.

However, considering that the method for generating the objective image was the same as the method for constructing it and the configuration of the method was too simple but the construction was not perfect, it can be suggested that the optimization problem caused in applying the constructive method to the objective image was hard for the used real-coded GA to precisely solve. If we apply the constructive method to a more complicated image, the constructive method might not be able to construct the complicated image satisfactorily.

The reason why it was hard for the used real-coded GA to solve the optimization problem might come from how to calculate the fitness values of the GA individuals. This time we calculated the fitness value based on the similarity between values on the corresponding pixels in the objective image and the constructed image. In this case, if there is a dominant color widely distributed in an objective image, any optimization methods should first try to adjust the parameters values of the method for constructing the objective image in order for a constructed image to have the similar distribution of the dominant color. As a result, any optimization methods are unlikely to adjust the parameters values corresponding to colors narrowly distributed in the objective image. We must pay attention to design of an optimization method including design of a way to calculate the fitness value when developing a new constructive method in future.

An another thing that we must pay attention to is that what structures a prepared method for constructing an objective image can and can not generate. In this paper, since we used the same method for both the generation and the construction of the objective image, we did not have to consider that. In fact, if an objective structure is not one of all the images that a prepared method for constructing an objective image can generate, it is impossible for the prepared method to construct the objective image. Therefore we have to reflect somewhat analysis of an objective image in design of a constructive method especially a method for constructing an objective image.

The concept proposed in this paper allows us to utilize constructive methods used in the past for a current task, so that it should be efficient to develop a new constructive method based on analysis of an objective structure while utilizing the past knowledge and experience for a current task. If we use interactions among components of a whole structure to generate portions of the whole structure, though we did not use them, it should be hard to obtain interactions rules based only on the analysis of an objective structure, and utilizing the past knowledge and experience should be helpful.

V. Conclusion

We proposed the concept of the constructive approach to understanding static structures. The concept is that we first obtain an interpretable process of constructing a given static structure using some model and then understand the given structure by means of the interpretable process of constructing it. To realize this concept, we need to combine a method for generating structures which provides an interpretable process of generating a structure with an optimization method which optimizes parameters values of the method for generating structures. In this paper, we combined the method for generating structures inspired by biological development with the real-coded genetic algorithm as an optimization method. The mechanism that enabled us to interpret a process of constructing a given structure was that the method for generating structures used in this paper made the initial structure representing an outline of a final structure gradually complicated toward a final structure while outputting medium structures.

We applied the realized constructive method to image understanding. We used the same method to both generate and construct the objective image, so that it was possible for the realized constructive method to perfectly construct the objective image if its optimization method could successfully adjust the parameters values of the method for generating structures. The results of the application showed that the construction of the given image was not perfect but the constructed images were similar to the given image in terms of not only a structure itself but also a process of construction. In this paper, since we connected similarity between processes of generating and constructing an objective image to possibility of understanding the given image, we could show the prospects of the proposed concept.

In the propose concept, how to design a method for generating structures and an optimization method to adjust parameters values of a method for generating structures is important. Designing those well might be more difficult than analyzing a given structure and extracting knowledge from it through its analysis. However, since the proposed concept allows us to utilize past knowledge and experiences on constructing static structure that seem useful for a current task, we do not always have to design the two methods necessary for crystallizing the proposed concept from scratch. In order to develop and improve a constructive method for understanding static structures, we need to newly design a method in parallel with accumulating cases of constructing a variety of structures with realized methods from now on. Furthermore, we can also expect to gain novel methods for generating structures through efforts of developing and improving constructive methods.

Acknowledgment

This work is sponsored by Ministry of Public Management, Home Affairs, Post and Telecommunications.

References

[1] B. Alberts and et al., "Molecular Biology of the Cell," Kyouikusya, 1990 (in Japanese).

[2] L. J. Eshelman and D. J. Shaffer, "Real-coded Genetic Algorithms and Interval-schemata," in D. L. Whiteley, editor, Foundations of Genetic Algorithms 2, pp. 187–202, Morgan Kaufman, San Mateo, CA, 1993.

[3] H. Kagawa, M. Kinouchi and M. Hagiwara, "Image segmentation by artificial life approach using autonomous agents," 1999 International Joint Conference on Neural Networks, CD-ROM, 1999.

[4] S. Kauffman, "At Home in the Universe: The Search for Laws of Self-Organization and Complexity," Oxford University Press, 1996.

[5] H. Kitano, "Systems Biology: A Brief Overview," Science, vol. 295, pp.1662–1664, 2002.

[6] C. G. Langton, "Artificial Life I," Addison Wesley, 1988.

[7] M. Gell-Mann, "The Quark and the Jaguar: Adventures in the Simple and the Complex," Henry Holt & Company, 1995.

[8] D. E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning," Addison-Wesley, 1989.

[9] D. E. Goldberg, "The Design of Innovation: Lessons from and for Competent Genetic Algorithms", Kluwer Academic Publishers, 2002.

[10] "Artificial Life: A Report from the Frontier Where Computers Meet Biology," Vintage Books, USA, 1993.

[11] K. Ohnishi and H. Takagi, "Feed-Back Model Inspired by Biological Development to Hierarchically Design Complex Structure," Proc. of IEEE Int'l Conf. on System, Man, and Cybernetics, pp. 3699–3704, 2000.

[12] H. Satoh and M. Yamamura and S. Kobayashi, "Minimal Generation Gap Model for GAs Considering Both Exploration and Exploitation," Proceedings of the International Conference on Fuzzy Systems, Neural Networks and Soft Computing (Iizuka'96), pp. 494–497, 1996.

[13] H. Takagi, "Interactive Evolutionary Computation: Fusion of the Capabilities of EC Optimization and Human Evaluation," Proceedings of the IEEE, vol.89, no.9, pp.1275–1296, 2001.

[14] H. Takagi and T. Noda, "Media Converter with Impression Preservation Using a Neuro-Genetic Approach," Hybrid Intelligent Systems, vol.1, no.1, pp. 49–56, 2004.

[15] S. Wolfram, "Cellular Automata and Complexity: Collected Papers," Addison-Wesley, 1994.

[16] K. Yoshida and T. Kato and T. Yanaru, "A Study of Database System with KANSEI Information," Proceedings of 1999 IEEE international Conference on System, Man, and Cybernetics, pp. 253–256, 1999.