Modeling and Evaluating the Process of Creating Paintings with Evolutionary Computing

Tomohiro Suzuki and Yasuhiro Suzuki
School of Information and Sciences, Nagoya University
ysuzuki@nagoya-u.jp

Abstract

We have modeled the actual process of creating paintings using Evolutionary Computing. In general, an outline and a theme are decided at the beginning of creating a painting, and events are located by the theme, however we evolve initial images to the preselected theme image with this method. We conducted a questionnaire to make the images created by this method to evaluate, and researched which generation people were interested in. We found that the images that have high aesthetic evaluation and originality are created in early and middle generations.

1. Introduction

The evolutionary algorithm has been employed for optimum solution searching and machine learning in the field of Evolutionary Computing, and it has achieved great results. However, there are still difficulties in applying these techniques to the art field because it is pretty difficult to conceptualize creative artifacts by solving a problem using genetic algorithm or optimization with multi-agent method.

The study in this field started with the Biomorph created by R.Dawkins (Dawkins, 1986), and there are two main types of approaches. One is Interactive Evolutionary Computing (IEC) and the other is Evolutionary Computing (EC). The former is a technique for optimization based on the user’s subjective evaluation, in simple terms, the technique replaces the fitness function by human beings.

IEC method is heavily used for evaluating paintings because it is difficult to numerically express the optimization performance as an evaluation function. However it fatigues users during evolution and has many problems. For example, it cannot have many generation changes, the speed of evolution is slow, and it can’t create many descendants due to the limited window space, and so on.

For these reasons, we employed EC method in this study and developed a method that simulated the real drawing process. In the drawing process, the composition and theme is decided first, and the layout for the motifs follows the theme. By employing EC, a user can set an initial image before the start of evolution, and then he can evolve it into the image he wants to output. This method has achieved rapid generation change, unlike IEC, because it only requires users to choose an initial image without fatigue. We conducted a questionnaire for evaluating the output images using Semantic Differential, SD method to detect which generation’s images are more creative. The questionnaire results show that high aesthetic and novel images were created in early and middle generations.

2. Back Ground

2.1 Previous research using IEC method

The Biomorph in The Blind Watchmaker by R.Dawkins is the first paper on the application to paintings with Evolutionary Computing method. It is easy to create various biological bodies like insects by repeating subjective preferences and mutations using L-system (Lindenmayer, 1968) which mathematically expresses the process of self recursive development of plants and applying the number of branches and angle parameters as genetic information. (Fig. 1)

Fig. 1 Example of output images of the Biomorph

From then on, a wide variety of studies have been conducted to output paintings using many IEC methods such as morphogenesis of insects, genesis of plants based on L-system and creation of 2D graphic arts based on mathematical formulas and cellular automata. Sims (Sims, 1991) who is known as a pioneer of IEC graphic artists has evolved mathematical formulas and created images the user wants by using GP techniques and evaluating CG images corresponding to the mathematical formula. (Fig. 2). And Secretan has created genetic arts using a kind of neural networks, Composition Pattern-Producing Networks that is n-
dimensional function in n-dimensional space and using IEC method called the Picbreeder (Secretan, 2011).

Fig. 2 Example of output images by Sims

2.2 Previous research using EC method
Most of studies on the creation of computer graphics using adaptive evolution have employed the foregoing IEC method, and there are not many studies that have used the EC method or the natural selection based on inter-component and environmental interactions. Among these studies, we focused on McCormack (McCormack, et al. 2012) who employed the process of niche formulation, and showed that each agent drew various paintings under regulated conditions (Fig. 3).

Fig. 3 Example of output images by J. McCormack

Unemi (Unemi, 1999) has created a system that locates some self-propagating individuals in two-dimensional Euclidean space and prevents them from colliding with each other during growth, and generates complex color patterns by coordinating a separate color for each individual. In addition to this, Bird et al. (Bird et al. 2008) attempted to create a robot that could draw lines as a drawing robot system. It doesn’t limit the robot’s drawing, but expresses “Ecological model” and is made up of the interrelationship between “Information from the surrounding environments” and “The way to reflect it to the robot’s drawing”. Also, Nakayama et al. (Nakayama, et al. 2012) have developed a new method that automatically creates the image in the user’s mind by making cut-and-paste processing of an existing picture which has textures and colors that the user imagines into a drawing operator (Fig. 4).

Fig. 4 Example of the evolution of paintings

The users will be able to perform the fitness calculation and the output of various paintings preferred by the individual users will be expected. However, IEC method is faced with some problems at present. First, generation change takes so much time because the user has to decide the evolutionary direction by generations. Second, it is difficult to evolve with many individuals because the pictures should be displayed on the screen by generations. Third, the user requires a certain level of training for selecting pictures, and it is especially-pronounces in abstract images that are difficult to be scored. Takagi (Takagi, 2001) has developed a technique for limiting the search range of evolution as much as possible and reducing the user’s repetitive selection in order to deal with such problems. Also Auerbach (Auerbach, 2012) has invented a method of automatic evolution by entering an image output from the Picbreeder into Google image search, and giving the high fitness to the frequently searched images.

However, there are still problems in determining the fitness for painting evolution on the study with EC method. As in the method above, for example, the most searched image is far from the picture that satisfies viewers. Also studies using both methods ignore the theme determination for painting. Usually a theme is firstly decided when a picture is painted by hand with painting tools. The creator starts to decide what information he wants to convey. Then, the layout of compositions and events, and coloring will be arranged in accordance with the theme. From this point of view, we consider that the IEC method is unfit for deciding a theme because the user needs to selected new image with respect to each generation, so the EC method which the user can firstly decide the initial image is fit to model the actual drawing process.

In this study, we consider “the theme of paintings” that has hardly been discussed and develop the model closely similar to the actual drawing process. In addition, we create images that provide the user inspiration for making a picture, and verify that the evolutionary computing technique is applicable to creative artifacts. Also we carried out a questionnaire survey to evaluate images output from the model.

3. Model

3.1. Model outline
We construct a system that automatically outputs paintings by use of genetic algorithm. First we get the user to select some favorite images by feeling and to decide one image as a theme.
Individual piece is composed of cut-and-paste pictures and collaged images, and the fitness is given to each piece according to how much the piece is similar to the theme image. The individual piece, which is given the high fitness, can leave more images of itself to the next generation, and the piece with the low fitness is more likely to be eliminated by natural selection. Also some images randomly change by mutation.

There are three reasons that we employed genetic algorithm. First, it is easy to evolve images for many generations and it evolves images more rapidly than the IEC method. Second, the number of individual pieces in each generation is not limited, so it is possible to evolve images with a lot of individuals. Third, the user can initially decide the target image as the theme, so he can expect the exact evolution toward the theme in his mind compared to the IEC method whose images keep on evolving impromptu. Usually, the theme is decided at the beginning of drawing a picture, and then the layout of events, composition determination, coloring, and balancing in accordance with the theme are conducted. We thought such a structure that moves towards one theme or optimum solution is easy to simulate by genetic algorithm.

3.2.1 Method for vertical and lateral partitioning
We describe the method of vertical and lateral partitioning. A filial image is created by combining one half each of selected parents’ images. If the parents’ images are cut in half lengthwise, combine a half of each image in a horizontal direction, or if they are cut in half crosswise, combine a half of each image in a vertical direction and random numbers decides the cutting position (Fig. 5).

![Fig. 5](https://example.com/fig5)

3.2.2 Collage method
We describe the collage method. It means to paste the selected mother’s image that is enlarged or reduced, translated, and rotated in the selected father’s image. (Fig. 6 above) A random number determines how much the mother’s image is deformed and where it is pasted. Affine transformation is used for changing images, and it is commonly used to deform images because it facilitate image processing by representing image scaling, translation, and rotation.

![Fig. 6](https://example.com/fig6)

In this study, we set the maximum magnification as 1.5 to prevent the original image from getting out of shape because of being enlarged too much.

3.2.3. Etching method
We describe the etching method. Paste the selected mother’s image after the affine transformation in the selected father’s image. If a pixel of the mother’s image is black and also the father’s one is black, change the father’s one to white. In case that the mother’s one is black and the father’s one is white, leave the father’s one as white. After this process is applied to all the pixels, the father’s image can appear from the mother’s image as a white picture. (Fig 6. below)

3.3. Mutation
After crossing over, some filial images mutate by the random numbers. Determine each pixel on the filial image by the threshold. If a pixel is below the threshold, it is reversed from black to white or from white to black (Fig. 7). The user can set the threshold at the beginning of the model. If the threshold is set low, the mutation is likely to occur more easily.

![Fig. 7](https://example.com/fig7)

3.4. Elite preservation strategy
The population of the next generation is stochastically formed in the selection of parent images based on the fitness of each individual. For this reason, even though good individuals are created by crossing over or mutation, sometimes they can’t survive to the next generation. To avoid this consequence, a
method was proposed to leave individuals that have high fitness in the population of individuals to the next generation. We refer to this as Elite Preservation Strategy.

In this model, the user can arbitrarily decide whether or not to conserve elites in the evolution. Also the user can initially decide the number of the elites conserved and the total number of individuals in each generation. The more the number of individuals, the more the probability of diverse evolutions, but it takes longer to compute, so it requires attention.

3.5. Method for evaluating the fitness
We describe the method for evaluating the fitness. A created filial image is given points according to the degree of coincidence with a target image. The target image is classified into the following three areas in accordance with the brightness value.

Area 1: If the filial image coincides with the target image, it is given D_1 point. (If it doesn’t, D_2 points are deducted.);
Area 2: If the filial image coincides with the target image, D_3 points are deducted. (If it doesn’t, it is given D_4 points.);
Area 3: There is no change whether or not the filial image coincides with the target image.

However, the Area 3 can get points if the user arbitrarily input points. When the total number of matched pixels in Area 1 is n_1, mismatched pixels is n_2, the total number of matched pixels in Area 2 is n_3, mismatched pixels is n_4, and then the fitness F of the filial image is F=D_1 n_1+D_2 n_2+D_3 n_3+D_4 n_4. The points that are given when it matches or mismatches are arbitrarily input at the beginning of the simulation.

![Image](image_url)

Fig. 8 Example of computing the fitness

We describe a specific example with Fig. 8 that shows a pattern diagram of a partial enlargement of an actual image. There are 4 match pixels and 1 mismatch pixel in red area, 1 match pixel and 6 mismatch pixels in blue area, and 1 match pixel and 3 mismatch pixels in green area between the filial image (a) and the target image (b). Therefore, the fitness F of the filial image is expressed as F=2·4+(-1)·1+(-2)·1+1·6, and then we obtain F=11.

4. Method for evaluating output pictures

4.1. Questionnaire using Semantic Differential method
We classified output pictures by generations such as early, middle, and late period, and then we did a questionnaire using Semantic Differential, SD method (Osgood, 1957) for researching what evaluation of output pictures were given in which generation. SD method is one of methods for evaluating sensibility aspects and it is used in many fields.

The feature of SD method is to evaluate a certain stimulating concept with a pair of adjectives. First, we draw up a questionnaire with pairs of adjectives that describe contrasting words and set a scale of 1 to 5 or 7 among a pair of adjectives. We evaluate a concept with respect to each pair of adjectives, and then we can compare the differences among stimuli. This method has been used for psychological experiment since early times, and deserves a long-established method. We used the studies by Tsutsui et al. (Tsutsui, et al. 2001) and Inoue et al. (Inoue et al. 1985) as a reference for pairs of adjectives.

4.2. Factor analysis method
Factor analysis is a method for specifying some elements that affect a number of observed data. This element is called a factor. In this study, we regarded the questionnaire items using SD method as observed data, and used this method for searching factors that affected those results. We conducted the factor analysis using Excel 2010 (Social Research Information Co., Ltd.).

4.2.1. Correlation coefficient
We can examine the strength of relationship between two items with correlation coefficient. If the correlation coefficient is positive, the item has positive correlation, and if it is negative, the item has negative correlation. The survey of the correlation coefficient of each item suggests how the items relate to each other. It is called a correlation matrix whose correlation coefficient is calculated from each item and the coefficient is described in a matrix.

4.2.2. Identification of the number of factors
We conduct the factor analysis with the correlation matrix and determine the number of factors. The following standards are generally used for determining the number of factors.

1. Gutman standard: Employ a factor whose eigenvalue is more than 1.
2. Scree standard: Plot the size of the eigenvalue and observe the change, and then extract the factor.
3. Employ the factors those of cumulative contribution ratios are more than approximately 60 %.
4. Employ a factorial structure with possible meanings.

In this study, we decided the number of factors as 2 factors that had high contribution ratios to interpret the results easily.

4.2.3 Factor nomenclature
We conduct a factor analysis with the number of factors determined in the previous section. We denominate the factor by reference to the factor loading of the questionnaire item that is affected by the factor.

4.2. Multiple comparison test
Multiple comparison test is used for comparing each data group without raising the significance level when there are more than 3 data groups. In this study, we compared each item in early, middle, and later generation and examined the significant difference with Tukey-Kramer method.
5. Results

5.1. Output results
We explain the output results of the model with or without elite preservation.

5.1.1. Output results with elite preservation
The following is the output results with elite preservation. We used 4 pieces of initial images (Fig 5.1. (a)) and one target image (Fig 5.1. (b)).

The number of individuals in each generation n=80, the mutation rate mRate=0.000005, image size is 600×800 pixels. The score when the black area of the initial image matches that of the target image, D_1=10; the score when the black area of the initial image mismatches that of the target image, D_2=-5; the score when the blue area of the initial image matches that of the target image, D_3=-10 and the score when the blue area of the initial image mismatched that of the target image, D_4=5. (Fig 9. (c)) shows that the output results of individuals that had the highest fitness in each generation and (Fig 9. (d)) shows that the fitness in each generation.

5.1.2. Output results without elite preservation
The following is the output results without elite preservation. We used initial images (Fig 9. (a)) and one target image (Fig 9. (b)).

The number of preserved elite individuals in each generation is 5. (Fig 10. (a)) shows the output results of individuals, and (Fig 10. (b)) shows the fitness in each generation.

If an individual with the high fitness is generated by crossing-over or mutation, it is not passed to the next generation because the elite preservation isn’t conducted. So we can see that the individual repeats evolution and selection. The feature of this case is that the initial image can’t get closer to the target image but a dynamic image tends to be output. Comparing to the case with elite preservation, the picture evolution isn’t fixed even though the generations roll by and an interesting picture can be output in later generations.
5.2. The results of the questionnaire
The following is the results of the questionnaire. The questionnaire was conducted at the booth of the creator’s market vol.2914 in Port Messe Nagoya from 7th to 8th December 2013. The creator’s market is a big event that attracts more than approximately 3000 people including professionals and amateurs for selling all sorts of original works such as fashions, interiors, crafts, visual designs, etc. We joined this event because we expected to get answers in a wide age group and have an opportunity to show a number of people output images. We got answers from 19 male and female participants between teenagers and those in their 50’s in this event. The results of the questionnaire are shown as points of SD method. We prepared three sets of generated images in early, middle, and later stage (Fig. 11). Each image was printed on a paper, and they were arranged in random order.

Fig. 11 Images used for the questionnaire
In the questionnaire results, high values are observed in early and middle generation, especially dynamic motion in middle generation shows extremely high values (Fig. 12). The three items “pleasure”, “fun”, and “originality” in early generation exceed those in middle generation, but the item “beauty” is at the same level. All the items in later generation are below those in early and middle generation, and especially dynamic motion is low.
5.2.1. Factor analysis
We describe the results of the factor analysis of the questionnaire. We got the cumulative contribution ratio and eigenvalue plot from the factor analysis without deciding the number of factors. The number of factor fields is numbered in descending order of the contribution ratio for descriptive purpose. If the number of factors is determined using the Gutmann standard and Scree standard, the number of factors becomes excessive, so it is difficult to grasp the structures. Therefore, we employ two factors that their eigenvalues are more than 3 and have extremely high contribution ratios in order to facilitate the interpretation. We conducted the factor analysis of these two factors using the maximum-likelihood method and pro matrix method. We found that the first factor had high factor loadings related to aesthetic evaluation of the output picture such as pleasure and beauty in early and middle generation, and then we named this factor as “aesthetic evaluation” factor.

<table>
<thead>
<tr>
<th>(Early)</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comfortable</td>
<td>0.5965</td>
<td>0.3899</td>
</tr>
<tr>
<td>Beauty</td>
<td>0.5411</td>
<td>0.4548</td>
</tr>
<tr>
<td>Dynamics</td>
<td>0.1918</td>
<td>0.4923</td>
</tr>
<tr>
<td>Pleasure</td>
<td>0.382</td>
<td>0.5761</td>
</tr>
<tr>
<td>Fun</td>
<td>0.1238</td>
<td>0.7894</td>
</tr>
<tr>
<td>Originality</td>
<td>0.0904</td>
<td>0.8835</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Middle)</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comfortable</td>
<td>0.7263</td>
<td>-0.0682</td>
</tr>
<tr>
<td>Beauty</td>
<td>0.7626</td>
<td>-0.1485</td>
</tr>
<tr>
<td>Dynamics</td>
<td>-0.162</td>
<td>0.4064</td>
</tr>
<tr>
<td>Pleasure</td>
<td>0.6504</td>
<td>-0.2494</td>
</tr>
<tr>
<td>Fun</td>
<td>0.8558</td>
<td>0.0162</td>
</tr>
<tr>
<td>Originality</td>
<td>0.7779</td>
<td>0.0534</td>
</tr>
</tbody>
</table>

Table 1. Results of factor analysis
The second factor had extremely high factor loadings related to fun and originality in early generation, the second highest factor loading was dynamics in early, middle, and later generation, and fun in later generation except for pleasure and beauty. These are related to originality and movement of pictures, so we named this factor as “originality” factor (Table 1.).

5.2.2. Multiple range test
We conducted multiple range test on each item in the results of the questionnaire and examined the significant difference (Table 2.). The results show 1% significant difference between early and later generation on pleasure items, between early and middle, and between middle and later generation on dynamic items, and between early and later generation on originality items. Similarly, there were 5% significant differences between early and later generation on beauty items, between middle and later generation on favorability items, between early and middle generation on fun items, and between middle and later generation on originality items.

<table>
<thead>
<tr>
<th>(e): Early, (m): Middle, (l): Later</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(e) Comfortable (m) Comfortable</td>
<td>0.2733</td>
</tr>
<tr>
<td>(e) Comfortable (l) Comfortable</td>
<td>0.0014  **</td>
</tr>
<tr>
<td>(m) Comfortable (l) Comfortable</td>
<td>0.0885</td>
</tr>
<tr>
<td>(e) Beauty (m) Beauty</td>
<td>0.9697</td>
</tr>
<tr>
<td>(e) Beauty (l) Beauty</td>
<td>0.0316  *</td>
</tr>
<tr>
<td>(m) Beauty (l) Beauty</td>
<td>0.0556</td>
</tr>
<tr>
<td>(e) Dynamics (m) Dynamics</td>
<td>0.0506  **</td>
</tr>
<tr>
<td>(e) Dynamics (l) Dynamics</td>
<td>0.3673</td>
</tr>
<tr>
<td>(m) Dynamics (l) Dynamics</td>
<td>0.0001  **</td>
</tr>
<tr>
<td>(e) Pleasure (m) Pleasure</td>
<td>0.8577</td>
</tr>
<tr>
<td>(e) Pleasure (l) Pleasure</td>
<td>0.1284</td>
</tr>
<tr>
<td>(m) Pleasure (l) Pleasure</td>
<td>0.0394  *</td>
</tr>
<tr>
<td>(e) Fun (m) Fun</td>
<td>0.8228</td>
</tr>
<tr>
<td>(e) Fun (l) Fun</td>
<td>0.0085  **</td>
</tr>
<tr>
<td>(m) Fun (l) Fun</td>
<td>0.0401  *</td>
</tr>
</tbody>
</table>
| (e) Originality (m) Originality    | 0.8044  *
| (e) Originality (l) Originality    | 0.0025  **|
| (m) Originality (l) Originality    | 0.0347  *|

Table 2. Results of the multiple range test

6. Discussion

6.1. Review of the results
The results of the questionnaire with elite preservation show that they came out of the opposite of the model fitness. From the table of actual factor analysis, it is believed that the fitness evenly becomes higher at every generation, on the other hand both originality and aesthetic evaluation are high in early generation, and then the originality comes down but the aesthetic evaluation goes up at the generations roll by. And the aesthetic evaluation is decreased in accordance with the later generation (Fig. 13).

These results show that interesting images are output in early generation whose fitness is low or in middle generation whose fitness starts to increase. Therefore the model that simply evolves and aims at the target image is inappropriate. This indicates that the model that sustains the fitness in early and middle generations can output better images. The models that stochastically conduct elite preservation or change points and constrain its evolution at the time of coincidence of images when the fitness becomes excess during evolution are good examples. The results show that the case without elite preservation keeps the fitness low, but a wider variety of output images are created than those in the case with elite preservation. When Fig. 14 is taken as an example, it evolves without elite preservation using a 4 leaf clover picture as a target image and 4 pieces of initial images that are used in the section 5, the output image becomes like a witch who looks to
the distance, so there is a possibility to output a heuristic image from the target image beyond all imagination.

Fig. 13 Schematic diagram of the changes of evaluation by generations

![Schematic diagram of the changes of evaluation by generations](image13)

Fig. 14. Pictures based on an evolved image

These results suggest that when the fitness in this model is employed, the area where we can expect an image having high aesthetic evaluation, originality, and heuristic output (Fig. 15).

Fig. 15 Schematic diagram of each area

These are rough considerations, but we could show that extremely simple models in this study can output various pictures. These results suggest that it is worthwhile creating a model that passes through the above area and carrying out further research.

Conclusion

We have developed Natural Computing researches, and it is our challenge to understand nature as algorithm; how computing halts is a basic problem in computing nature. We believe where “aesthetics” plays a key role in computing in most of nature, halt of computing is not defined in advance, but when computing processes reach comfortable state, the process halts. What is comfort and beauty in Aesthetics has been considered, and it corresponds to “computational aesthetics” in Natural Computing context, but computational aesthetics is different from traditional Aesthetics or philosophy, comfort and beauty are considered from a viewpoint of algorithm. We have been interested in verbal feature and computing of tactile sense, and developed computing aesthetics on tactile sense proposing tactile scores for describing tactile sense as well as the pictures and their impressions in this study.

Acknowledgements

JSPS KAKENHI Grant Number 90217513 and 50177044 supported this work.

References


C. E. Osgood, G.H. Suci and P.H. Tannenbaum, (1957), The measurement of meaning, University of Illinois presses.


Masaki Inoue and Toshinobu Kobayashi, (1985) The research domain context, but computing in advance, so there is a possibility to output a heuristic image from the target image beyond all imagination. The results suggest that when the fitness in this model is employed, the area where we can expect an image having high aesthetic evaluation, originality, and heuristic output (Fig. 15).

Fig. 15 Schematic diagram of each area

These are rough considerations, but we could show that extremely simple models in this study can output various pictures. These results suggest that it is worthwhile creating a model that passes through the above area and carrying out further research.