Visualization of EC Landscape to Accelerate EC Conversion and Evaluation of its Effect

Boschetti, Fabio
Commonwealth Scientific and Industrial Research Organization

Takagi, Hideyuki
Faculty of Design, Kyushu University

https://hdl.handle.net/2324/1808439

Visualization of EC Landscape to Accelerate EC Conversion and Evaluation of its Effect

Fabio Boschetti

Commonwealth Scientific and Industrial Research Organization

39 Fairway, Nedlands 6009, Western Australia

fabio@ned.dem.csiro.au

http://www.ned.dem.csiro.au/BoschettiFabio

Hideyuki Takagi

Kyushu Institute of Design

Shiobaru, Minami-ku, Fukuoka, 815-8540 Japan

takagi@kyushu-id.ac.jp

http://www.kyushu-id.ac.jp/~takagi

***Abstract*—** **We evaluate how visualization of an evolutionary computation (EC) landscape is effective using a geophysical task. This technique allows us to actively participate in EC optimization by viewing the distribution of searching points on 2-D space mapped from an n-D EC landscape, and indicating where in the EC is the possible global optimum. We construct a Visualized GA system that includes self-organizing maps for visualization and compare its performance with that of a normal GA using the geophysical simulation task. Sign tests for the comparisons show that the Visualized GA converges significantly faster than the normal EC ($p < 0.01$), which suggests further extensions to enhance user interactivity.**

1 Introduction

Interactive evolutionary computation (IEC) is an optimization approach where the EC optimizes target systems based on human evaluation. Simply speaking, the IEC is an EC whose fitness function is taken place by a human user.

The IEC has been applied to several tasks in artistic, engineering, edutainment, and other fields. Some of them are CG design, melody and rhythm generation, industrial design, speech and image processing, hearing aid fitting, data mining, virtual reality, media database retrieval, robotics, and so on [9, 10, 11].

Since the IEC is an interactive optimization method between a human and a tireless computer, it has strong constraints on an EC search; both population size and maximum number of generations are restricted to 10 to 20 due to human fatigue. To solve this problem, several improvements to the human interface have been proposed [9, 10, 11]. A visualized IEC that allows IEC users to view the EC landscape and point out the possible global optimum is one of the proposals [12, 4].

The objective of this paper is to evaluate the effect of the visualization using not only a benchmark function but also a simple realistic task, a geophysical simulation.

We review the Visualized IEC or Visualized EC in section 2 and then evaluate the effect of the visualization of an EC search landscape using the geophysical simulation task in section 4. In this paper, we adopt genetic algorithms (GA) as one of the EC techniques. Also, we adopt Visualized GA instead of Visualized IGA to purely evaluate the effect of visualization.

2 Visualization of EC Search Landscape
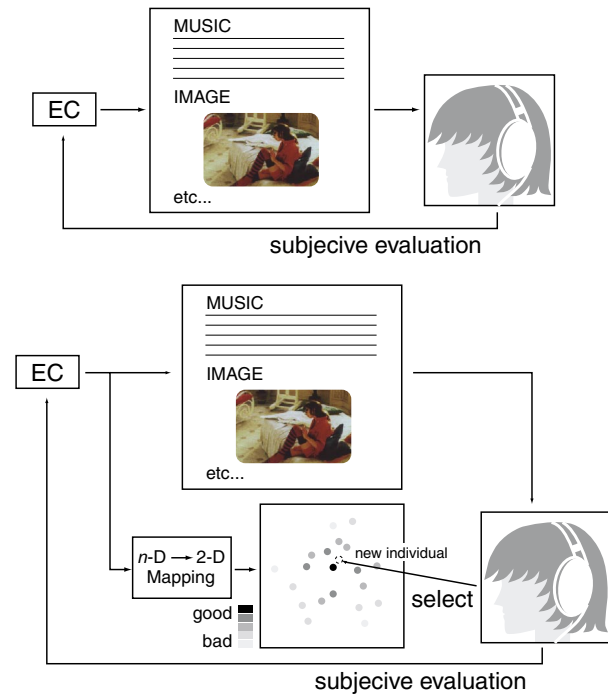
2.1 Visualized IEC / EC systems

Figure 1: Diagrams of IEC (upper) and Visualized IEC (lower).

The feature of Visualized IEC or Visualized EC is that the EC landscape of n-D space is mapped to 2-D space and the distribution of searching points is visualized by a human user [12, 4]. Since the human user is able to grasp rough landscape information from the distribution of searching points, he or she

can actively cooperate with the EC search by pointing out the possible global optimum point as an excellent individual choice. As the EC directly and systematically searches the original n-D space using EC operators, and humans grasp an entire distribution of individuals in the 2-D space at a macroscopic level, we can expect that the combination of these different capabilities results in a faster EC convergence.

Figure 1 shows a normal IEC and a Visualized IEC, and Figure 2 shows an example of Visualized GA.

Mapping from n-D space to 2-D space is an essential operation for the visualization. There are several mapping methods for 2-D visualization, for example, the principle component analysis, Sammon's non-linear mapping [8], self-organizing maps [5], Visor [6], TOPAS [7], and the method using Genetic Programming [13]. Any visualization methods are usable in the Visualized IEC, but some methods are more suitable for the Visualized IEC than others.

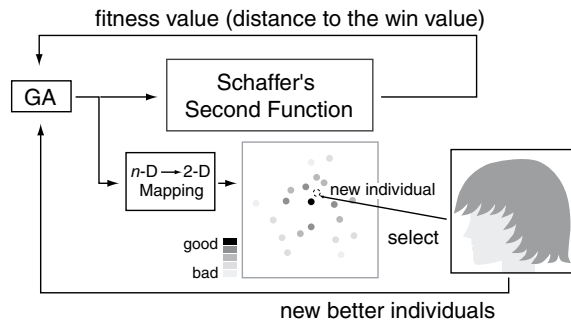## 2.2 Evaluation of Visualization Using a Benchmark Function

Figure 2: Experimental system of the Visualized GA. The GA determines the coordinate of the minimum value of the Schaffer's second function, and the difference of the function output and the minimum value is fed back into the GA as a fitness value. The human operator visually selects a possible global optimum in the mapped 2-D space and sends it to the GA as a new possible parent. A self-organizing map is used to map individuals from an n-D space to a 2-D space.

We constructed a Visualized GA system shown in Figure 2 and evaluated it; see its detail in reference [12, 4]. The reason for using this system instead of Visualized IGA is to avoid the fluctuations due to human subjective evaluation for all individuals, and to evaluate only the effect of visualization. The experimental system adopts self-organizing maps (SOM) among many mapping methods for the visualization.
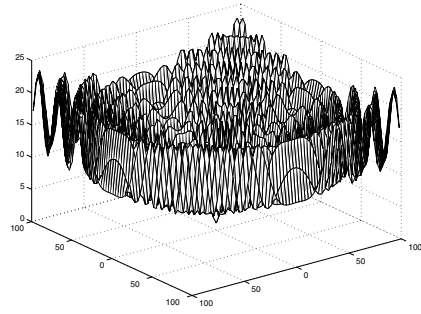

Figure 3: Modified Schaffer's second function given by $\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{4}} * \left\{\sin^2(50 * \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{10}} + 1\right\}$, where $-100 \leq x_i \leq 100$ and $n = 3$ and 5.

The experimental Visualized GA system has a population size of 20, and we compare it with a normal GA with population sizes of 20, 100, and 1000; these two systems correspond to the whole system and the upper loop only in Figure 2, respectively. The evaluation task uses the modified Schaffer's functions of 3-D and 5-D shown in Figure 3.

Figure 4 shows the experimental results. Solid lines are the results of Visualized GA systems conducted by five human subjects, and dotted lines are the results of normal GAs with population sizes of 20, 100, and 1000. These results imply that the performance of a Visualized IGA corresponds to 5 to 50 times of that of a normal GA.

3 Geophysical Optimization Task

In exploration geophysics, we attempt to reconstruct the distribution of rocks below the surface of the earth from measurements of their physical responses. This, in turn, is used to build a geological model that is able to suggest to expert geologists the possible location of mineral or oil resources.

Among the most traditional techniques, seismic data, gravity, electric, and magnetic measurements are collected, and the distribution of physical properties underground (like acoustic impedance, density, magnetic susceptibility) is deduced through an inverse process. This basically consists of a numerical optimization, in which the set(s) of geological parameters is searched for the best fit to the measured data, within the noise level.

Geophysical optimization problems are among the hardest; they are highly non-linear, of high dimensionality, often very computationally intensive, and highly non-unique. Geophysical optimization requests not only numerical simulation but also the expert's knowledge to interpret the numerical solutions and disregard unrealistic ones; numerical solutions without the interaction of experts have lit-
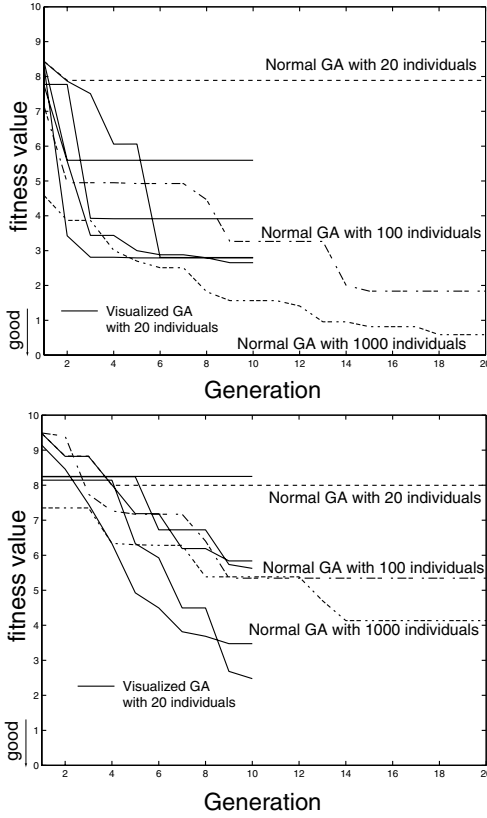
Figure 4: Convergence characteristics of normal GAs with population sizes of 20, 100, and 1000, and the Visualized GA with a population size of 20 for a modified Schaffer's function of 3-D and 5-D, are shown in the upper and lower graphs, respectively. Solid lines represent the five experimental subjects for the Visualized GA.

tle geological meaning in most cases. Consequently, tools are required that provide easy interactivity to geophysicists as well as less computational cost.

In this paper, we apply the Visualized GA used in section 2 to the inversion of gravity data and evaluate its effectiveness. Our aim is to reconstruct the location and shape of rocks from their gravity responses.

Figure 5 shows our experimental task: a synthetic vertical section through the shallow layers of the earth. The top of the section is the surface of the earth. The bottom represents the depth of 500 meters. The section is 12.8 km long.

The white rectangles represent rock bodies of anomalously high density. Figure 6 shows the gravitational response due to this model. High values of the gravity response lie above the high density bodies. The purpose of the experiment is to reconstruct the section in Figure 5 from the data in Figure 6. Clearly, this is an inverse process that can be cast
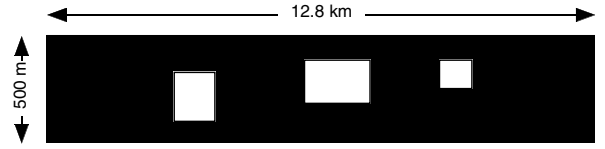

Figure 5: Synthetic vertical cross section, from the surface towards the center of the earth, that expresses a 500-meter depth and 12.8 km width. The white blocks represent rock bodies of high density.
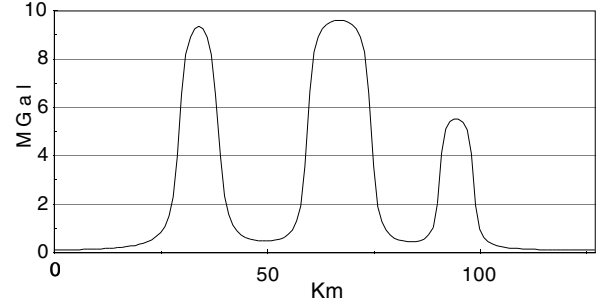

Figure 6: Gravity response of the model in Figure 5 at each 100 meters; x-axis corresponds 0 to 12.8 km.

into an optimization problem, where we search for the optimal parameters that describe the location and characteristics of the blocks to fit the gravity data. In particular, we assume *a priori* to know the number of blocks (this is geological information readily available: roughly one block for each bump in the gravity profile) and that the blocks have a rectangular shape (a reasonable geological approximation). We attempt to reconstruct the parameters describing the location, shape and density contrast for each block.

4 Evaluation of Visualization Effect using Geophysical Task

4.1 Experimental Setting

We compare a normal GA and a Visualized GA that adopts SOM, as for the system in section 2, using a geophysical task. The GA used in this section is a real-coded GA; see reference [1] for the details of the GA condition and a geophysical application.

We run two tests. In the first test, we aim to reconstruct the location and shape of each rock body, while assuming a known single density contrast. Fixing the density contrast of the bodies makes the inverse problem unique, and simplifies the interpretation of the results. See reference [3] for more information about uniqueness issues related to the inversion of gravity data.

In the second test, we adopt a different density

Table 1: Eight comparisons between normal GA and Visualized GA. The misfit values that are Euclidian distances between the measured data and estimated data produced by the best GA individuals for eight runs are shown here. Notice that the Visualized GA outperforms the normal GA in all runs and this difference is signified by a sign test with ($p < 0.01$). Also, notice that in all but the first run, the Visualized GA outperformed the normal GA in less than 20 generations; the number of generations by which the Visualized GA outperforms the normal GA is shown in brackets.

normal GA	Visualized GA
.21	.20
.32	.29 (in 10 gen.)
.22	.18 (in 13 gen.)
.21	.20 (in 14 gen.)
.27	.23 (in 11 gen.)
.31	.23 (in 9 gen.)
.36	.34 (in 13 gen.)
.50	.42 (in 8 gen.)

contrast for each rock. Although this condition does not make the problem exactly non-unique (in order to achieve strict non-uniqueness, we need to allow the shape of the bodies to vary also), we can now find models that are able to reproduce the data with a very low misfit, to the point that we can consider the problem non-unique within the resolution allowed by the GA process (also within the resolution allowed by realistic levels of noise in real data).

The misfit data in Tables 1 and 2 are the Euclidian distances between the measured data and estimated data produced by the best GA individual.

4.2 Results of the First Evaluation

Only the location and shape of the rectangular bodies are searched in this first evaluation. Each body is described by 4 parameters, the locations of its vertices; we optimize 12 parameters in total. Eight runs were performed with different initial GA seeds. Same seeds are used for a Visualized GA and some for a normal GA.

The population size is 10 and the number of GA run generations is 20. The reason we have to adopt such a small population size is that geophysical applications usually require such a long computation time that we we want to minimize the evaluation time.

The best misfit rate of the Visualized GA outperforms that of the normal GA in each of the eight trials, as shown in Table 1. Also, the Visualized GA needs less than 20 generations to outperform the normal GA except in one of the eight trials (see the

number of generations required in brackets). This better convergence of the Visualized GAs results in less computation time.

The best results obtained by the Visualized GA and the normal GA are shown in the upper and lower parts of Figure 7, respectively. Although the quality of the reconstruction seems comparable, the data misfit rate, that is the only measure of model *goodness* available to the algorithm, shows that the Visualized GA better optimizes the experimental task than the normal GA.

We should also notice that, despite the two solutions being comparable in final geological appearance, they have been obtained at very different computational costs: 20 generations for the normal GA and 13 for Visualized GA. This seems to suggest that the visualization of a GA landscape is most valuable when the function evaluation involves a considerable computational effort.
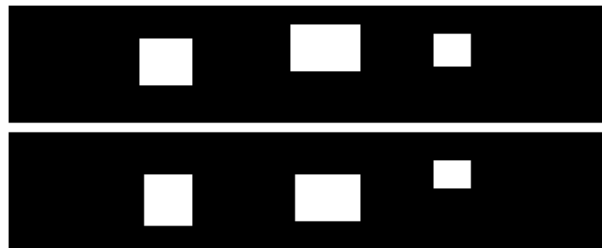

Figure 7: Best solution from the Visualized GA (upper) and from the normal GA (lower). The quality of the solutions seems comparable, but the one from the Visualized GA has been obtained at a much lower computational effort.

Finally, we show the samples of 2-D mapped individuals in Figure 8 for the reader's interest.

It is interesting to notice how the best individual, and individuals with fitness very close to the best (orange points in original color images), are occasionally generated by the GA process in unexplored areas, and how the GA population follows them in the following generations, also thanks to the user interaction.

4.3 Results of the Second Evaluation

In the second test, we do not assume that each rock density contrast is the same. The number of parameters to be optimized becomes 15: 12 parameters mentioned in the first evaluation plus 3 rock density contrasts. This also makes this task effectively non-unique. The population size and GA generations are same as those in the the first evaluation: population size of 20, and 10 generations. Although the number of trials in the second evaluation was only two,

Table 2: Comparison between normal GA and Visualized GA for the second evaluation.

normal GA	Visualized GA
.45	.36 (in 6 gen.)
.46	.44 (in 8 gen.)

the Visualized GA outperformed the normal GA in both cases. Table 2 shows the results of numerical comparison.

The best solutions found by the GA and Visualized GA are shown in Figure 9. The depth of gray color in the blocks reflects density contrast with the surroundings. From the physics point of view, we know that a low density body at shallow depth produces a similar response to that of a higher density, but smaller body, at greater depth. This is what happens for the Visualized GA solution (top) for the right-most body, and for the normal GA solution for the central body. In general the Visualized GA solution is closer to the synthetic model we aimed at reproducing.
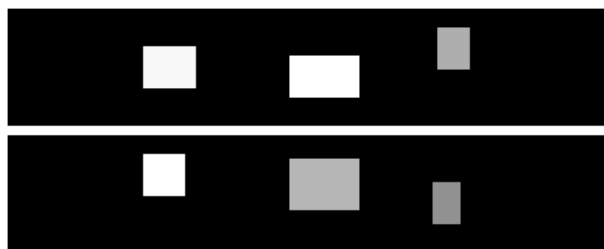

Figure 9: Best solution from the Visualized GA (upper) and from the normal GA (lower) obtained in the second test. The depth of gray color in the blocks reflects the degree of rock density contrast. The deeper gray color becomes, the lighter density contrast becomes.

5   Discussion and Future Directions

Most geoscientific problems are non-unique. This is often due to the fact that data are collected in 1-D or 2-D, while the source signal obtained from rocks is always 3-D. Other geoscientific problems may be characterized by the change of phases for geochemical problems or the sharp changes of behavior for geomechanical problems, for example, which depend on several control parameters.

Rough visual appraisal of the solution space is particular useful to understand these problems. It is important, for non-unique problems, to be aware of the non-uniqueness and its extent as well as finding one solution for these problems. Geoscientists need to know if different geological scenarios may be hidden behind the same data set. It is crucial, for problems characterized by different phases or mechanical behavior, to know their transition states. When this information cannot be obtained analytically, visual tools become extremely useful.

We have already applied normal IEC to optimize a geological problem via subjective evolution [2]. Although we evaluated the Visualized GA to highlight the visualization effect in this paper, we have used IGA (Interactive GA) so that the user more actively interferes with the optimization in our research. Subsequent to our previous geological IEC application, tedious manual analysis of the GA population was carried out to extract information about the search space and possible domains of different geological behavior. Visualized IEC must be an easy alternative approach for this purpose. We will explore this approach in our next experiments.

Finally, the Visualized EC or Visualized IEC can also be useful to deal with non-unique problems. As mentioned before, most geoscientific problems are mathematically non-unique. However, most solutions are often geologically *unrealistic* and can be dismissed from the point of view of geophysical knowledge and experience. The visual appraisal of such solutions could allow the user to direct the search towards geologically realistic solution models. For example, the Visualized IEC or Visualized EC can be used as shown in Figure 10 to help geophysicists in their decision making; the visualization has the option to display corresponding images and let geophysicists select according not only to its numerical fitness but also its geological realism. Experiments in this direction will be carried out in the near future.

6   Conclusions

Accelerating EC convergence using the visualization of the fitness distribution helps in guessing the EC landscape. Active human intervention in the EC search has statistically been proven effective, using not only a benchmark function in our previous work, but now, in this paper, with a more realistic geophysical task in both 12-D and 15-D space with different EC implementations. The experiments also suggest extensions to enhance the interactivity even further, by allowing the user to view the solutions and direct the EC search.

References

[1] Boschetti, F., Dentith, M., and List, R.: "Inversion of seismic refraction data using Genetic Al-
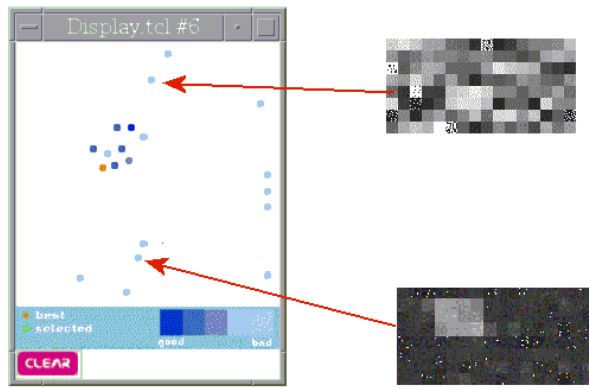
Figure 10: Possible enhancement for the visualization. Points in the display window could be visualized by the user in order to select them according to their geological feasibility. In this example, model (a) would be judged geologically *unrealistic*, while model (b) would be selected as *feasible*.

gorithms," Geophysics, vol.61, no.6, pp.1715–1727 (1996).

[2] Boschetti, F. and Moresi, L.: "Interactive Inversion in Geolsciences," Geophysics (in press).

[3] Blakely, R. J.: *Potential theory in gravity and magnetic applications*, Cambridge University Press, New York, NY, USA (1996).

[4] Hayashida, N. and Takagi, H.: "Visualized IEC: Interactive Evolutionary Computation with Multidimensional Data Visualization," IEEE Int'l Conf. on Industrial Electronics, Control and Instrumentation (IECON2000), Nagoya, Japan, pp.2738–2743 (Oct, 2000).

[5] Kohonen, T.: *Self-Organizing Maps*, Springer-Velag, Heidelberg, Germany (1995).

[6] König, A., Bulmahn, O., and Glesner, M.: "Systematic methods for multivariate data visualization and numerical assessment of class separability and overlap in automated visual industrial quality control," 5th British Machine Vision Conf., vol.1, pp.195-204 (Sept., 1994).

[7] König, A.: "A survey of multivariate data projection, visualization and interactive analysis," 5th Int'l Conf. on Soft Computing and Information/Intelligent Systems (IIZUKA'98), Iizuka, Japan: World Scientific, Singapore, pp.55-59 (Oct., 1998).

[8] Sammon, J. W.: "A nonlinear mapping for data structure analysis," IEEE Trans. on Computers C-18, no.5, pp.401-409 (1969).

[9] Takagi, H.: "Interactive Evolutionary Computation: System Optimization Based on Human Subjective Evaluation," IEEE Int'l Conf. on Intelligent Engineering Systems (INES'98), Vienna, Austria, pp.1–6 (Sept., 1998).

[10] Takagi, H.: "Interactive Evolutionary Computation – Cooperation of computational intelligence and human *KANSEI* —," 5th Int'l Conf. on Soft Computing and Information/Intelligent Systems (IIZUKA'98), Iizuka, Fukuoka, Japan: World Scientific, Singapore, pp.41–50 (Oct., 1998).

[11] Takagi, H., Unemi, T., and Terano, T.: "Perspective on interactive evolutionary computing," J. of Japan Society for Artificial Intelligence, vol.13, no.5, pp.692-703 (1998) (*in Japanese*).

[12] Takagi, H.: "Active User Intervention in an EC Search," Int'l Conf. on Information Sciences (JCIS2000), Atlantic City, NJ, USA, pp.995–998 (Feb. /Mar., 2000).

[13] Venturini, G., Slimane, M., Morin, F., and Asselin de Beauville, J. P.: "On using interactive genetic algorithms for knowledge discovery in databases," 7th Int'l Conf. on Genetic Algorithms, Morgan Kaufmann Publisher, pp.696-703 (1997).
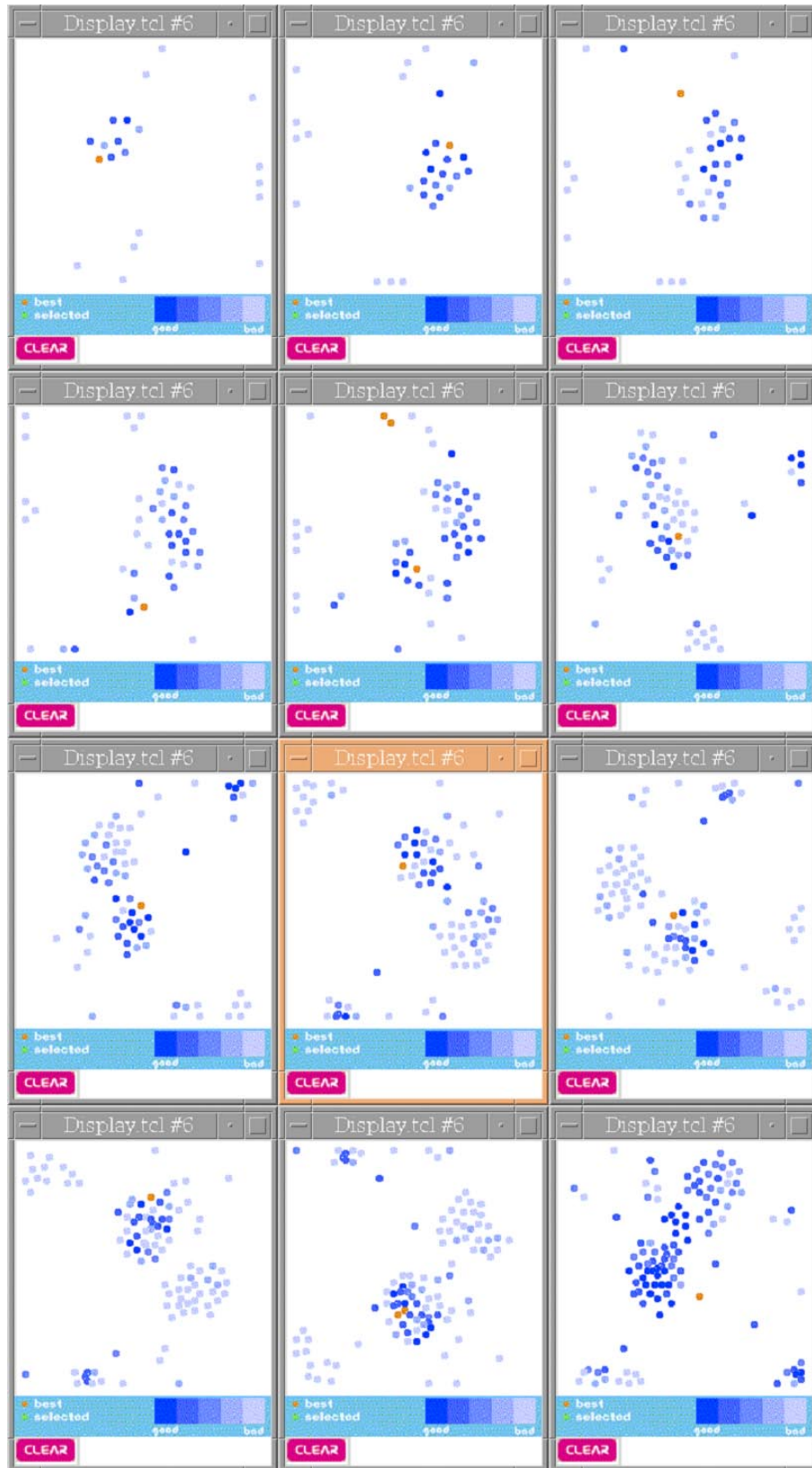
Figure 8: Visualization of the best GA run, from generation 2 to 13 (final). Original images have the five levels of depth of blue color for fitness value of each individual and orange color for the possible global optimum selected by a user.