H. Pohlheim

Advanced Techniques for the Visualization of Evolutionary Algorithms

1 Introduction

Evolutionary algorithms work in an algorithmically simple manner. When put to work they produce a vast amount of data. Apart from simple convergence information it is a non-trivial task to extract useful information from those data to provide insight into the state and progress of the evolutionary algorithm. Methods for extracting and visualizing relevant data are still under development ([9], [10]).

In this paper an advanced technique for visualizing multidimensional data produced by the evolutionary algorithm is presented. This advanced technique is independent of the representation of the variables. The technique is applied to the visualization of variables of individuals of a population. The use of this technique for the visualization of multi criteria objective values in multiobjective optimization is demonstrated. Especially useful is this technique for comparison of high-dimensional data of different runs.

Beside the advanced method a set of standard visualization techniques for different data types is presented. These data types and techniques give an instant visual impression of the evolutionary algorithms progress and the actual state of the individuals of the population. The methods were selected according to their usefulness for real world applications. Hopefully, the defined standard set will open up a discussion between designers and users of evolutionary algorithms about necessary and useful visualization techniques.

Section 2 describes multidimensional visualization and presents some examples. In Section 3 a standard set of visualization techniques is proposed. Section 4 gives concluding remarks and some prospect directions of further investigations.

2 Multidimensional Visualization

Most of the commonly used techniques for visualization are limited to representing data depending on one or two variables, because of our limitation for standard visualization to three dimensions. To break this limitation there exist two possible extensions: using color for the fourth dimension and time as the fifth dimension. Neither possibility is very common and requires practice, especially using time for visualizing the fifth dimension. However, if the problem incorporates more than five dimensions a new method for visualizing arbitrarily high dimensions must be found.

For the visualization of multidimensional data a method to transform multidimensional data to a lower dimension is looked for, preferably to 2 or 3 dimensions. This transformation should provide a lower-dimensional picture where the dissimilarities between the data points of the multidimensional domain correspond with the dissimilarities of the lower-dimensional domain. These transformation methods are often referred to as ‘multidimensional scaling’ ([2], [3], [8]).

To measure the dissimilarity, the distance between pairs of data points is used. These distances can be genuine distances in the respective high-dimensional domain, for instance the Euclidean distance. If a genuine distance measure is not applicable, the dissimilarities can be defined by a

substitute measure. The distances need not satisfy the triangle inequality[1]; thus the term dissimilarity is used. In [8] examples of non-metric measures of categorical data are given (for instance, simple matching coefficient).

2.1 Calculation of low-dimensional representation

For finding a low-dimensional representation, y_i, of a number of points in a high-dimensional domain, x_i, the distance between all points in their respective domains is calculated.

$x_1,\ldots,x_n$:	data points in $\Re^p$	δ_{ij}:	distance between x_i und x_j
$y_1,\ldots,y_n$:	data points in $\Re^q$, $p \geq q$	d_{ij}:	distance between y_i und y_j

Then a configuration of low-dimensional image points, y_i, is looked for, such that the distances d_{ij} between image points are as close as possible to the corresponding original distances δ_{ij}. Since it will usually not be possible to find a configuration for which $d_{ij} = \delta_{ij}$ for all i and j, a quality criterion, J_{ef}, for ranking different configurations must be defined, equation 1.

$$J_{ef} = \frac{1}{\sum_{i<j}\delta_{ij}} \sum_{i<j} \frac{\left(d_{ij} - \delta_{ij}\right)^2}{\delta_{ij}} \qquad \nabla_{y_k} J_{ef} = \frac{2}{\sum_{i<j}\delta_{ij}} \sum_{j \neq k} \left(\frac{d_{kj} - \delta_{kj}}{\delta_{kj}} \cdot \frac{y_k - y_j}{d_{kj}} \right) \tag{1}$$

An optimal configuration minimizes the criterion function and can be sought by a gradient descent method. The gradient of the criterion function is easy to compute, equation 1.

One of the best known methods for multidimensional scaling is SAMMON mapping [11]. SAMMON used a Newton method (steepest decent). However, this method is not very robust and diverges without special interaction [8].

For optimization a more robust search method was employed, the RPROP algorithm [7]. The RPROP algorithm uses just the change of the sign of the gradient for step size control. RPROP is widely used in the area of neural networks.

The starting configuration can be chosen randomly. However, using a principal component analysis for defining the starting configuration, the convergence of the optimization can be accelerated. Nevertheless, because of the local search the optimization often gets stuck in local optima, thus multiple runs using different initializations were done. After comparing the results the best configuration was chosen.

A few papers concerning the use of multidimensional scaling for the visualization of evolutionary algorithm were published recently ([1] and [4]).

2.2 Visualization examples of multidimensional scaling

This section presents visualization examples that use multidimensional scaling to map data to a low-dimensional domain for visualization.

- visualization examples of individuals of a population and
- visualization examples of non-dominated solutions in multiobjective optimization.

2.2.1 Variables of best individual of every generation

This example shows the „path through search space“ of the best individual of every generation. The resulting graphic produces a better understandable picture than visualizing the best individual of all generations. By multidimensional scaling the variables of the best individuals of all

[1] ‘triangle inequality’: $d_{ik} \leq d_{ij} + d_{jk}$, d: distance between data points.

generations were reduced to two dimensions. These low-dimensional data were visualized employing a 2-D point plot. The label of every point is the generation of the respective individual.

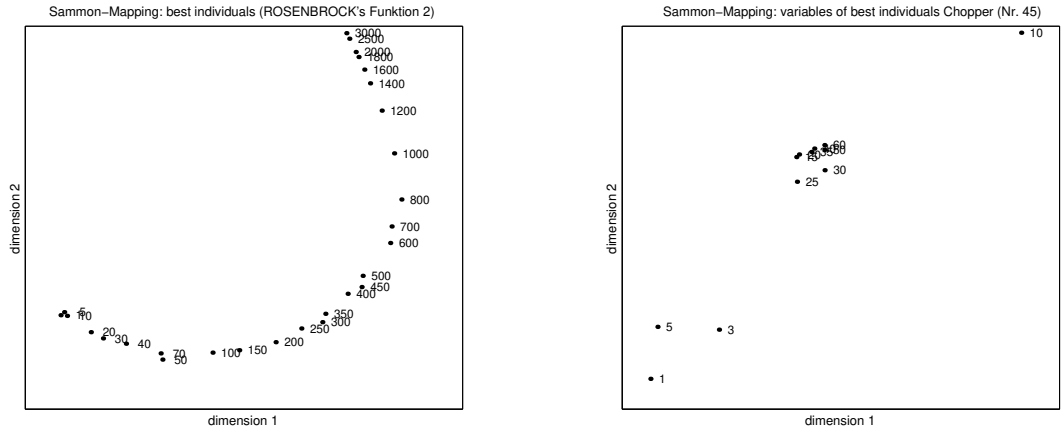

Figure 1: Multidimensional scaling of variables of best individuals of one run; left: ROSENBROCKs function, right: Chopper system

Figure 1 presents the results of two different objective functions. The left plot is a run of ROSENBROCKs function, often called ‘banana function’. The used function employed 10 dimensions. The optimization ran over 3000 generations. The plot clearly shows the curve-shaped characteristic of this function known from 2-D plots of this function.

The second example, figure 1 right, shows the result of a real world application, the optimization of a Chopper (DC/DC line-side converter), where one individual consisted of 9 variables here. The characteristic of the path of the best individual through search space is quite different compared to ROSENBROCKs function. During the first 5 generations the best individuals are in one area. In generation 10 the best individual is quite far away. Through all further generations the best individuals can be found in a very small area.

2.2.2 Multiple objective values of best individual of every generation

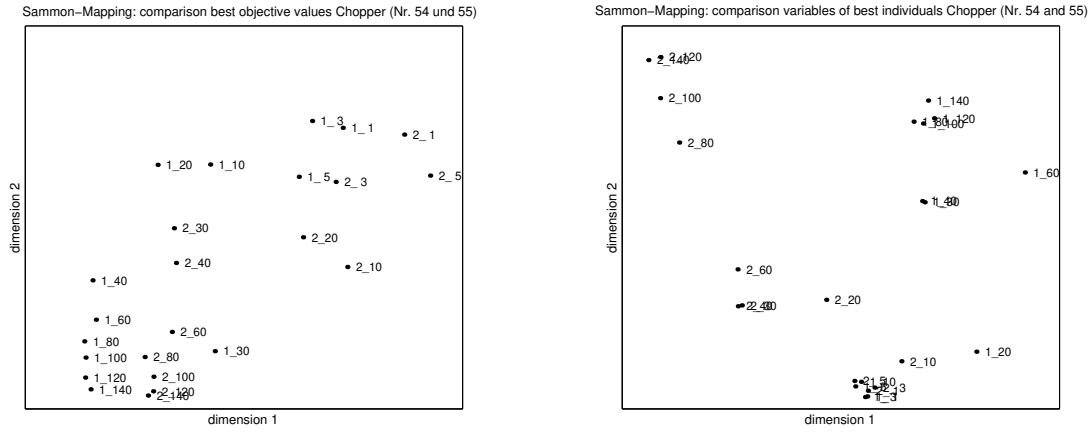

Figure 2: Comparison of two optimization runs of the Chopper application using multidimensional scaling; left: multi criteria objective values, right: variables of corresponding individuals

The second example presents the multidimensional scaling of multi criteria objective values. Again we use the real world Chopper example. The objective function consisted of the simulation of up to nine different scenarios, where each scenario produced one objective value. The

sum of all separate objective values was used as the objective value for the evolutionary algorithm. However, it proved difficult to compare the course of two different runs, but was solved using multidimensional scaling.

Figure 2 shows the comparison of two optimization runs of the Chopper. On the left side the visualization of the multi criteria objective values is shown, on the right the best individuals of each generation. The label of each point consists of the number of the run followed by the number of the generation. The graphics show that both runs start from the same area of the objective space or variable search space. However, during the run the objective values and the individuals both diverge to different areas along a line each. Without multidimensional scaling it would be very difficult to extract this information from the results.

3 Set of standard visualization techniques

In this chapter a number of data types and techniques useful for visualizing the state of the population and the progress of the evolutionary algorithm are proposed. These techniques are simple to calculate and easy to display using standard visualization tools. This set of techniques provides a base line for understanding the evolutionary process.

The techniques are systematized into two categories depending on the time range the data represent. The first category encloses methods for visualizing data produced during many generations, thus presenting a picture of the progress of the evolutionary algorithm. The second category contains techniques to visualize data produced anew for every generation. These techniques give a picture of the current state of the population or the evolutionary algorithm.

For each data type proposed a number of techniques is specified and visual examples are given. The information to be extracted from the respective graphics is briefly described.

3.1 Visualization of the course of the evolutionary algorithm

The following data types are used for visualizing the progress of the evolutionary algorithm:

1. objective value of best individual of every generation (convergence diagram),
2. variables of best individual of every generation,
3. objective values of all individuals of every generation.

3.1.1 Best objective value (convergence diagram)

The visualization of the objective value of the best individual of every generation is used very often, figure 3 left. Depending on the scaling of the objective values this graphic may give a good impression of the convergence of the evolutionary algorithm.

Many publications also include the mean and worst objective value in the convergence diagram. However, including the worst objective value very often destroys the scaling of the graphic.

More information can be extracted, when the mean objective value and the standard derivation of all objective values is included instead, figure 3 middle. This proofed appropriate for many real world problems solved using evolutionary algorithms.

When the regional model (often called migration model or coarse grained model) is employed the simple convergence plot can be extended to show the best objective value of each subpopulation, figure 3 right. This is of further advantage when different strategies for every subpopulation are used.

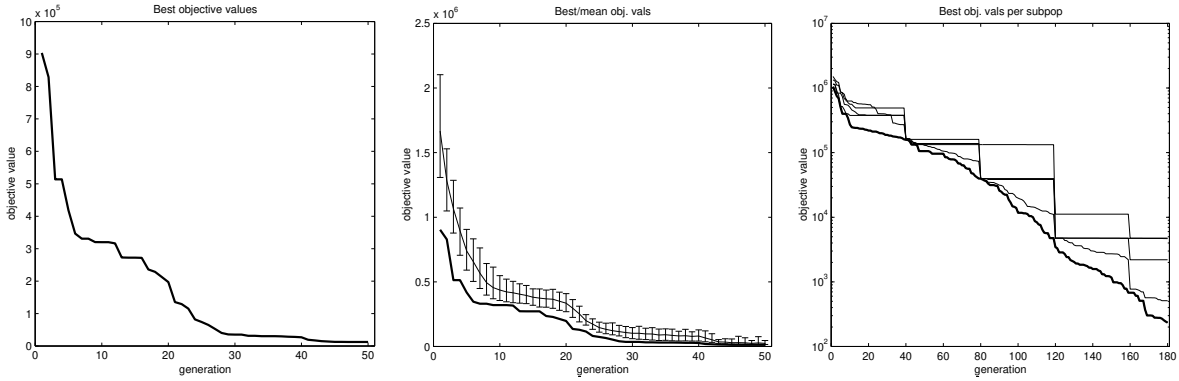

Figure 3: Convergence diagram; left: objective value of best individual, middle: best and mean objective value including standard derivation of all objective values, right: best objective value of every subpopulation

3.1.2 Variables of best individual

The visualization of the variables of the best individual of every generation is one of the most representative graphics for the progress of an optimization run. The user can actually see how the variables are changed and which combinations of variables are successful. During the run the user gets a picture of good variable values.

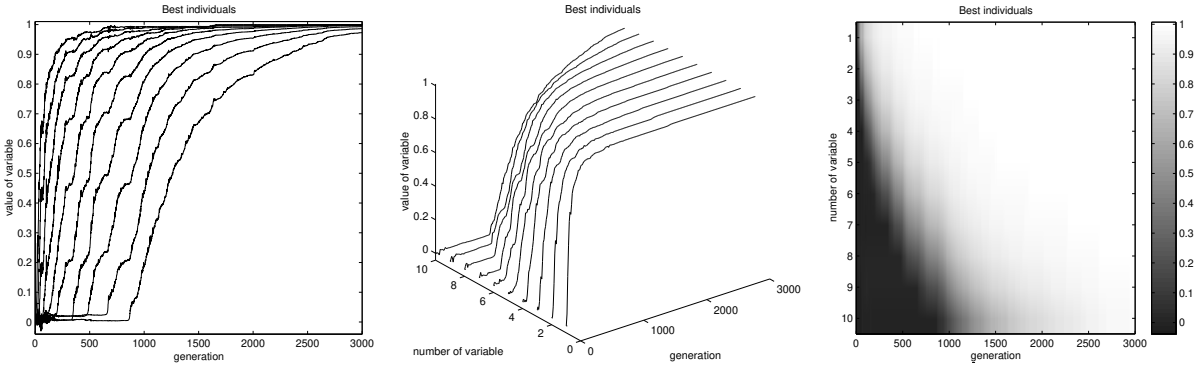

Figure 4: Variables of best individual of many generations; left: 2-D line plot; middle: 3-D line plot, right: 2-D image plot

Figure 4 presents three different techniques for this data type. The simplest one is the 2-D line plot, figure 4 left. Every line represents the course of one variable. However, the assignment between variable number and line is not possible (without a legend). The 3-D line plot, figure 4 middle, provides the extra dimension to assign every variable visually to a line. Nevertheless, the best technique to extract all this information is a 2-D image plot, figure 4 right. Even for a high number of variables and different values of adjacent variables a good impression of the distribution and change of the variables is possible.

Very different ranges of the variables can be circumvented by scaling the variables according to the domain of each variable, thus visualizing the relative values of the variables.

3.1.3 Objective values of all individuals

The previous two data types involved just the best individual of the population. To get an overview of the whole population the objective value of all individuals of every generation can be visualized.

Two techniques are demonstrated. The 2-D point plot, figure 5 left, shows the distribution of the objective values. The 2-D image plot, figure 5 middle and right, additionally shows the assign-

ment between objective value and position of the individual in the population. This is especially useful when the regional model, figure 5 middle, or the local model, figure 5 right, is employed. In figure 5 middle the changing size of the subpopulations due to competing subpopulations is apparent.

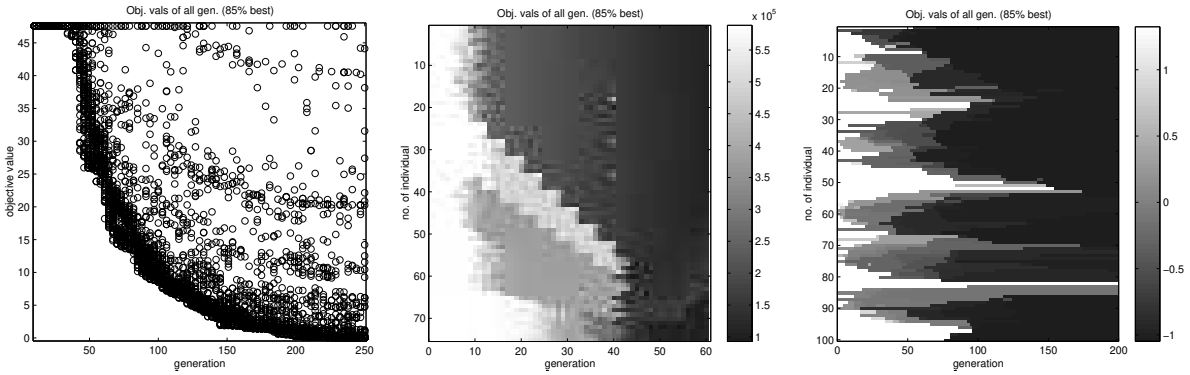

Figure 5: Objective values of all individuals of many generations; left: 2-D point plot, middle: 2-D image plot employing a regional model with competing subpopulations, right: 2-D image plot employing a one-dimensional local model

3.2 Visualization of the state of the evolutionary algorithm

The following data types and techniques are proposed for visualizing the actual state of the evolutionary algorithm:

1. objective value of all individuals,
2. variables of all individuals,
3. distance map of objective values and distance map of variable values.

3.2.1 Objective values of all individuals

Similar to the variables of the individuals the objective values of all individuals of one generation give a good impression of the state of the population.

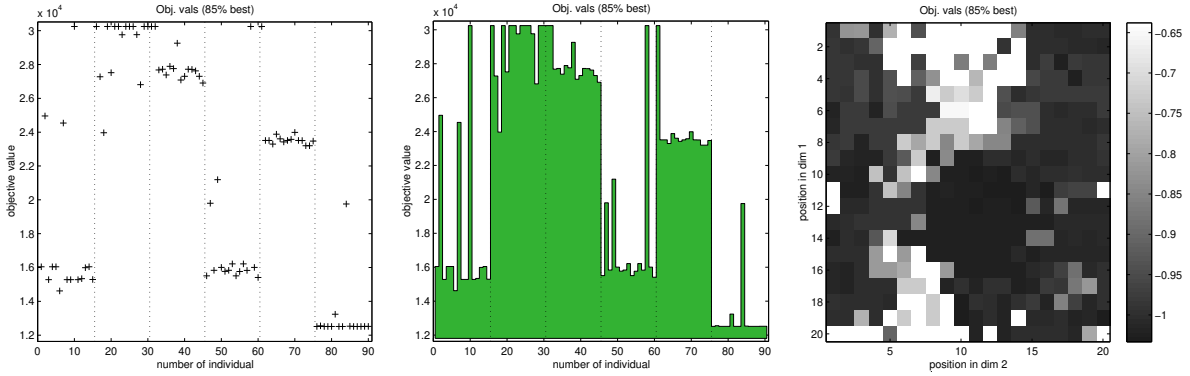

Figure 6: Objective values of all individuals of a population; left: 2-D point plot, middle: 2-D stairs plot, right: 2-D image plot employing a two-dimensional local model

Figure 6 presents three techniques for visualizing the objective values. The 2-D point plot in figure 6 left is straightforward. By using a filled 2-D stairs plot, figure 6 middle, the differences inside the population and adjacent individuals are more apparent. Both graphics display a population divided into subpopulations, the dotted lines display this division. Figure 6 right employs a 2-D image plot to visualize the two-dimensional neighbourhood of the objective values when using a two-dimensional local model. A sequence of such graphics shows the diffusion of good objective values through the population during an optimization run.

3.2.2 Variables of all individuals

The visualization of the variables of the individuals of a population provides an insight into the distribution of the individuals, the distance between individuals and the concentration of the individuals in one or more areas. However, because of the sheer amount of information the chosen technique for visualization is important.

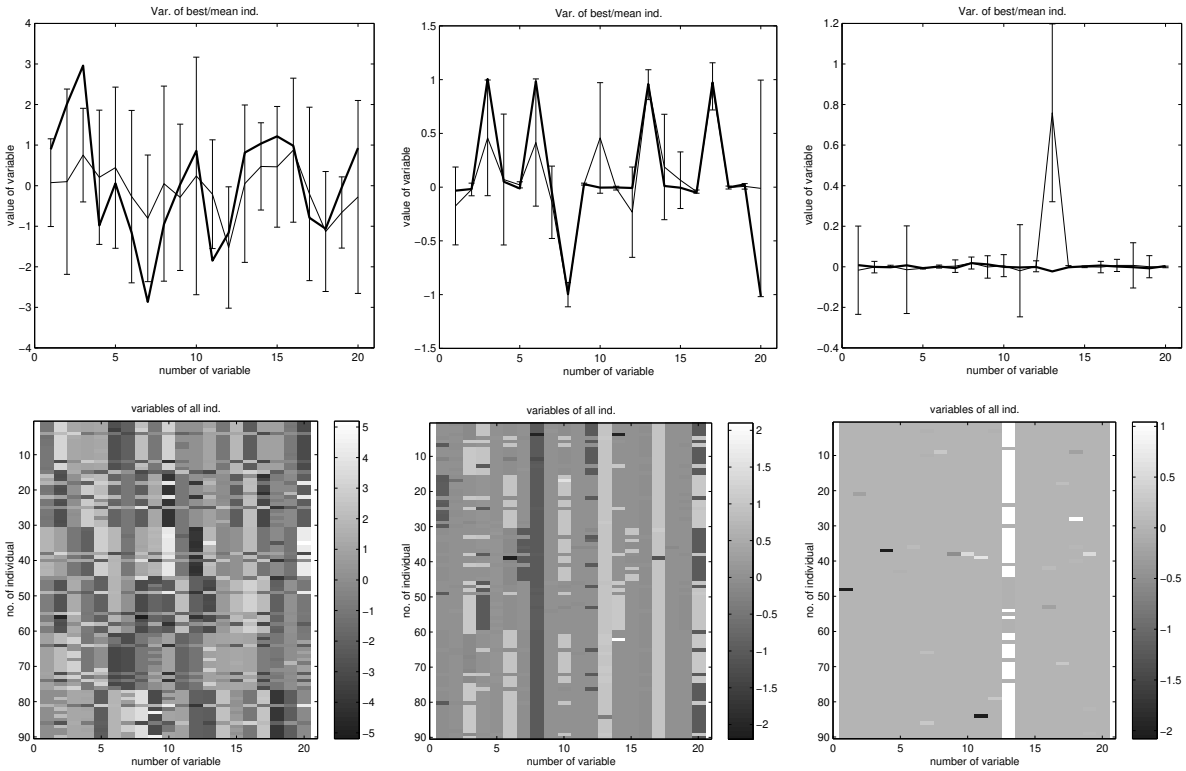

Figure 7: Variables of all individuals of a population; top: 2-D line plot of the best (thick line) and mean individual including standard variation of all variables, bottom: 2-D image plot; each row shows three stages of an optimization run - left: begin of run, middle: middle of run, right: end of run

Figure 7 shows two techniques each at three different stages of an optimization run. Figure 7 top presents the use of a 2-D line plot of the variables of the best individual together with the mean and the standard derivation of the variables. The sequence of three graphics displays the change of the variables values and their distribution. Figure 7 bottom employs a 2-D image plot to directly visualize the variable values. Due to the extra dimension more information can be extracted.

3.2.3 Distance between individuals

All previous data types were directly produced by the evolutionary algorithm. The distance between individuals, however, is a derived data type. For the calculation of the distance the Euclidean distance between the variables of the individuals is used.

If there is no neighbourhood relationship between individuals (global and regional model) the distance distribution is visualized using a 2-D histogram stairs plot, figure 8 left. This graphic directly shows the distribution of the individuals.

When the local model is used a neighbourhood relationship is defined. Thus, the distance should be visualized using the individuals in their respective neighbourhood. In a one-dimensional local model every individual has two neighbors, in a two-dimensional local model four neighbors. The distance between individuals of an one-dimensional local model is shown in figure 8 middle,

employing a 2-D stairs plot. The 2-D image plot in figure 8 right displays the distance between individuals in a two-dimensional local model.

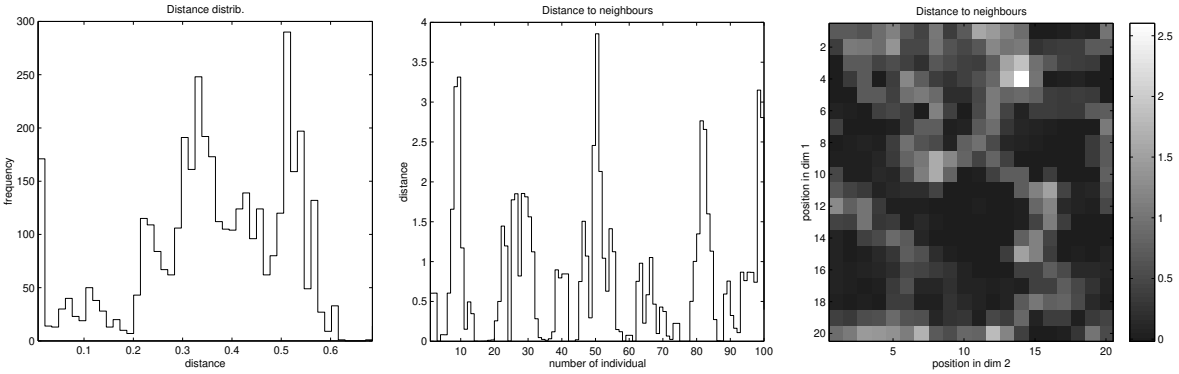

Figure 8: Distance between individuals of a population; left: distance distribution, middle: 2-D stairs plot of distance map employing a one-dimensional local model, right: 2-D image plot of distance map employing a two-dimensional local model

All data types and techniques visualizing the state of a population or the evolutionary algorithm should be used in a sequence of the respective graphics during an optimization run or afterwards. The animation of these graphics can reveal new and extended information and insights into the problem at hand.

4 Concluding remarks

This paper proposed multidimensional scaling for visualization of data produced by evolutionary algorithms. The examples presented showed the advantage of this technique. Information, which is not available or very difficult to extract using other techniques can be derived. Multidimensional scaling opens a field for new visualization techniques not only applicable to the domain of evolutionary algorithms.

The proposed set of "standard" visualization techniques of evolutionary algorithms provides a base line for understanding the evolutionary process. Combined with advanced techniques powerful visualization tools to aid the designer and user of evolutionary algorithms can be constructed.

The presented visualization techniques were applied in the course of solving real world problems using evolutionary algorithms. Thus, these techniques are not just proposed. These techniques were used and proved informative. An example implementation of all these and additional techniques are contained in [6].

References

[1] *Collins, T. D.*: Genotypic-Space Mapping: Population Visualization for Genetic Algorithms. personal communication, 1997.

[2] *Cox, T. F. and Cox, M. A. A.*: Multidimensional Scaling. London: Chapman & Hall, 1994.

[3] *Duda, R. O. and Hart, P. E.*: Pattern Classification and Scene Analysis. New York: John Wiley & Sons, 1973.

[4] *Dybowski, R., Collins, T. D. and Weller, P. D.*: Visualization of Binary String Convergence by Sammon Mapping. in *Fogel, D. B.*: Evolutionary Programming V, Proceedings of the Fifth Annual Conference on Evolutionary Programming. Cambridge, Massachusetts: MIT Press, 1996.

[5] *Fonseca, C. M. and Fleming P. J.*: On the Performance Assessment and Comparison of Stochastic Multiobjective Optimizers. in *Voigt, H.-M., Ebeling, W., Rechenberg, I. and Schwefel, H.-P.*: Parallel Problem Solving from Nature - PPSN IV: International Conference on Evolutionary Computation. volume 1141 of Lecture Notes in Computer Science, Berlin, Heidelberg, New York: Springer-Verlag, pp.584-593, 1996.

[6] *Pohlheim, H.*: Genetic Algorithm Toolbox for use with Matlab - Documentation. Technical Report, Technische Hochschule Ilmenau, 1996.

[7] *Riedmiller, M. and Braun, H.*: A direct adaptive method for faster backpropagation learning: The RPROP algorithm. in H. Ruspini (editor), Proceedings of the IEEE International Conference on Neural Networks (ICNN), pp.586-591, 1993.
[8] *Ripley, B. D.*: Pattern Recognition and Neural Networks. Cambridge, GB: Cambridge University Press, 1996.
[9] *Routen, T. W. and Collins, T. D.*: The Visualisation of AI Techniques. Proceedings of Third International Conference on Computational Graphics and Visualisation Techniques COMPUGRAPH'93, Alvor, Portugal, New York, USA: ACM Press, pp. 274-282, 1993
[10] *Routen, T. W.*: Techniques for the Visualisation of Genetic Algorithms. in Proceedings of The First IEEE Conference on Evolutionary Computation, Piscataway, New Jersey, USA: IEEE Service Center, Vol. II, pp. 846-851, 1994.
[11] *Sammon, J. W. jr.*: A Nonlinear Mapping for Data Structure Analysis. IEEE Transactions on Computers, vol. C-18, no. 5, pp. 401-409, 1969.

Author:

Pohlheim, H.
Technical University Ilmenau, Faculty of Informatics and Automation,
PSF 327, 98684 Ilmenau, Germany,
E-mail: pohlheim@systemtechnik.tu-ilmenau.de