

INTERACTIVE SONIFICATION MONITORING IN EVOLUTIONARY OPTIMIZATION

Florian Grond¹, Oliver Kramer², Thomas Hermann¹

¹Ambient Intelligence, CITEC Bielefeld University, Universitaetsstraße 21 - 23 33615 Bielefeld Germany

²Bauhaus-Universität Weimar Institut für Strukturmechanik, Marienstraße 15 99423 Weimar, Germany

[fgrond, thermann]@techfak.uni-bielefeld.de

oliver.kramer@uni-weimar.de

ABSTRACT

This case study introduces interactive sonification to evolutionary strategies (ES) for global optimization. We briefly describe the specific strengths of sonification as a tool for monitoring, the emerging trend of interactive sonification, and what it can add to the field of evolutionary computation. Then we line out the background of ES as optimization heuristics, briefly explain the algorithmic procedure of ES and discuss the need to intervene during optimization runs and the current shortcomings in appropriate user feedback. This motivates the development of an auditory closed loop setup that brings the expertise of interactive sonification to the field of monitoring ES algorithms. Further, we present considerations for the sound design and the detailed mapping of parameters from the ES to sound properties. Finally, we discuss the various implemented modes of interaction and their significance for the optimization through ES.

1. MONITORING THROUGH SONIFICATION

In all the different fields of applications, monitoring is amongst the most suitable ones for the use of sonification. Well-established examples range from the operating theatre to monitoring seismograms by listening [1]. Sonification has further been used for the monitoring of stock markets [2], network traffic [3], electrocardiograms [4], quantum oscillations [5], and EEG data [6][7]. The widespread use as a monitoring tool is because the human auditory system is particularly apt for this task. The two most important listening abilities for the purpose of monitoring are backgrounding, which sets in when a sound becomes steady, as well as the ability to focus on selected streams in a mixture of sounds [8]. Additionally, the auditory system has a strong capacity to readily notice transient sounds. Finally, and most important for monitoring, the user does not need to have a particular orientation in space in order to follow a process by listening.

2. FIELD OF APPLICATION

In our work we apply interactive sonification to evolution strategies (ES), which have grown into powerful optimization heuristics [9]. ES algorithms are biologically inspired, population based, randomized search heuristics. ES apply the principles of biological evolution to optimization: inheritance and mutation of genes, and selection of the fittest solutions according to the famous Darwinian principle. In the sixties and seventies, Fogel [10], Holland [11], Rechenberg [12] and Schwefel [13] translated these paradigms into algorithms that are called evolutionary computation today, a field that has evolved and diversified into a rich and frequently used set of methods for optimization problems. The resulting algorithms search efficiently for optimal solutions in high dimensional parameter spaces.

2.1. The principles of ES

For a better understanding of the sound design, which will be described later, we briefly introduce in more detail the underlying principle of the optimization procedure. The aim of finding an optimal solution corresponds to finding the global minimum of a cost-function, which is usually embedded in a high dimensional parameter space. ES are particularly useful in black box optimization scenarios, e.g., when no derivatives are available, which could be invested into the search process. This is why ES contain a random element in exploring the search space.

- The initial step is to randomly select a set (population) of points (individuals) that covers the search space of interest.
- Secondly, their fitness, which corresponds to the value of the cost-function, is evaluated.
- In a third step a defined percentage of the species with the poorest fitness values are discarded (selection).
- Forth, offspring is produced by new species that are derived from the fittest by varying their position in parameter space with a certain mutation strength (distribution σ) around their ancestors (inheritance).

This procedure is repeated until the selection of points converges towards the optimal solution. The development of the cost function and the σ vector for examples of a converging and a non-converging optimization is shown in figure 1.

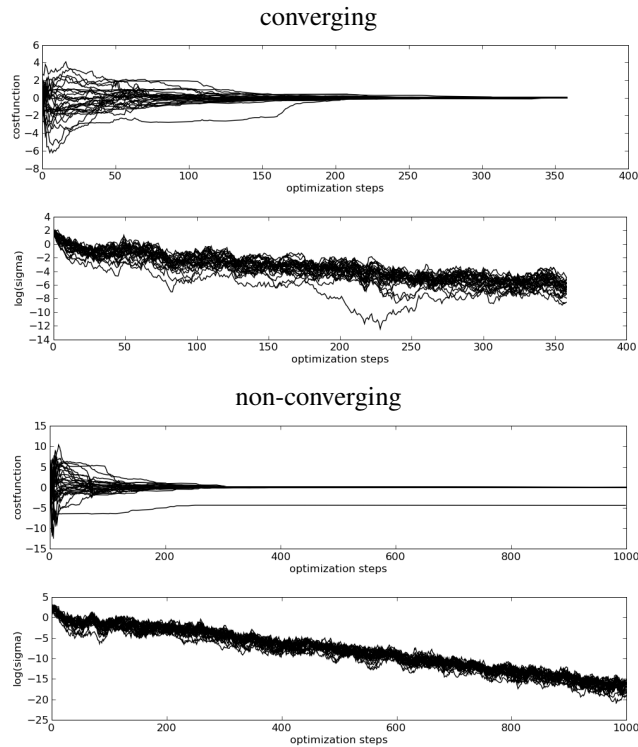


Figure 1: The converging optimization on top terminates after 359 iterations. The non converging optimization is interrupted after 1000 iterations. In the plot of the costfunction one component reaches a plateau which suggests that the optimization got stuck in a local minimum. In the depicted case the costfunction is known to have its global minimum at the origin, i.e 0 in all components.

2.2. Why Monitoring the Evolutionary Search Process?

In practice, the success of stochastic methods is fairly parameter dependent, and their tuning becomes one of the most important issues for successful optimization. A self-evident example is the control of mutation strengths that have to decrease as the population approaches the optimum in order to finally converge to the optimal solution. Automatic parameter tuning methods like sequential parameter optimization [14] for offline tuning or self-adaptation [15] [16] for online control of parameters are frequently applied.

Nevertheless, many practical problems still remain hard to solve even with the support of automatic parameter tun-

ing. Therefore the practitioner would appreciate the possibility of monitoring for potential real-time intervention that allows to control important parameters before and during the optimization process.

An essential prerequisite for this is that more knowledge about the search process could be gathered through appropriate feedback during the run. In this case user interventions would turn into a closed-loop interaction with the algorithm.

The current situation for evolutionary algorithms is that real time displays are restricted to quite poor interfaces from an HCI perspective. In most cases the operator monitors the state of the algorithm by watching a flow of output numbers. This is a cumbersome process, since many parameters change simultaneously at high speed. This fast changing information is in turn an ideal case to be displayed by sonification since the temporal resolution of our auditory system considerably exceeds the one of our vision.

3. THE APPLICATION DESCRIBING THE SOFTWARE COMPONENTS

In the following we describe the setup for the monitoring sonification of ES. An important prerequisite for the acceptance of novel displays is a good integration of already existing tools from the targeted fields into the application setup. As a widespread platform we rely on the programming language Python with the numerical extension NumPy, where extensive libraries and components for scientific computation are available. The sonification is implemented with the sound synthesis language SuperCollider3 (SC3) since it allows for versatile and advanced sound synthesis as discussed in [17]. For ES a good scalability on the sound synthesis side is particularly important, since the ES have scalable parameters, one of which would be the number of species. The communication between both parts is accomplished by the Open Sound Control protocol (OSC) [18]. This framework ensures a professional environment on both ends, the numerical computation as well as the sound synthesis.

3.1. Sound Design

3.2. Monitoring Processes and Auditory Augmentation

The challenge of the sound design is to create a monitoring display that maximize information and minimized at the same time intrusiveness, as discussed by Vickers in [19]. Vickers discussed three types of tasks for auditory process monitoring and their relation to two different types of receiving/perceiving information from an auditory display, namely hearing (PUSH) and listening (PULL). We list here the first two, which concern us mostly in the context of ES monitoring:

- Direct auditory display (PULL characteristic): the information to be monitored is the main focus of attention and does not allow for parallel activities.
- Peripheral auditory display (PUSH characteristic): attention is focused on a primary task whilst required information relating to another task or goal is presented on a peripheral display.

The new trend of auditory augmentation as presented by Bovermann et al. [20] allows for interesting hybrid solutions within the two suggested categories above. Whilst the interactive triggering of a sound would certainly be considered as a PULL activity, the sonic information can be in certain circumstances very tightly integrated into an everyday activity, which is very informative but not distracting.

One example of a concrete implementation that inspired us from [20], is the *Reim* framework, which augments the typing impact sounds on a keyboard. For augmentation usually a contact microphone is used. For our prototype we found that the inbuilt microphone on a MacBook Pro picks up the typing sounds good enough for proof of concept testing. We think that *Reim* is particularly useful for our context, since monitoring ES during multitasking on a computer work place means interacting with other programs at the same time. This means that the sound scape of the workplace is not polluted through additional sounds. However, the state of the optimization can be deliberately *queried* by just scratching onto the keyboard casing. These two aspects of non-distracting integration yet interactive triggering of a sound leads to interactive monitoring. Two activities that seem to mutually exclude each other at first.

The augmentation of the impact sounds is usually achieved through filtering. For a better recognizability of different states of progress during the optimization, we make use of filter stacks, where each individual filter is adjusted according to the actual solution in the parameter space of the cost-function.

3.3. Mapping the ES optimization to sound

The general goal of the sonification was to display the progress of the optimization. More specifically, the sonic information should also contain cues about the state within the parameter space and the progress of convergence:

- First, the sonification should include hints about the mutation strength for each dimension of the actual solution.
- Second, the progress of the optimization should be noticeable on all scales, starting with the strong variations at the beginning, but should also display small adjustments during the stage of convergence towards the end of the optimization.

- Third, the sonification should help to discern different areas within the parameter space such that convergence to different local minima can be acoustically distinguished.

For the sound design we oriented our choice based on what the ES algorithm suggested as usable metaphors. The most important value to be mapped to sound is the space of the cost-function, with no a priori range limit. For typical optimization test case like (e.g. Schwefel's or Griewank's multimodal function, [21]) the parameters in the solution space of interest are positive real numbers with ranges of different magnitudes. The second important feature from ES is the mutation-strength which we mapped to filter widths as an apt metaphor.

3.3.1. Mapping the Space of the Cost Function

Each dimension of the cost-function was represented through a stack of 5 band pass filters, with decreasing level $[0.0, -6.0, -9.5, -12.0, -14.0]$ dB. The centre frequencies of the filters were integer multiples of a base frequency f_b , that was unique for each dimension n . In order to span the whole audible frequency range, the base frequencies were linearly and equidistantly distributed between the MIDI notes 15 and 80. This means that the lowest f_b was 19.445 Hz and the highest frequency from the filter-stack with the highest $f_b = 830$ Hz was 4153 Hz. The f_b s of each filter stack were multiplied with a factor that corresponded to the magnitude of the component from the current solution of the cost function.

For a better distinguishability, the filter stacks were equidistantly distributed across the stereo panorama, alternating left and right with respect to the increasing base frequencies. The result was a unique set of spatially distributed bandpass filters for each point of the parameter space. In difficult optimization problems a different optimum can be found for each run, which means in turn that a different timbre would be heard at the end of each optimization. The timbres of various points from parameter space have been systematically sampled and were found to be noticeably distinct for most cases. All sound can be accessed here ¹.

3.3.2. Mapping the Mutation Strength

During an optimization run, the values of the cost-function typically converge very fast at the beginning with an exponential decay towards the optimum. After this fast transient moment it usually takes a while for the species to settle around and come close to the optimum. During this phase the change of the cost-function values is minimal, but this is often the crucial moment when the algorithm gets stuck

¹<http://www.techfak.uni-bielefeld.de/ags/ami/publications/GKH2011-ISM>

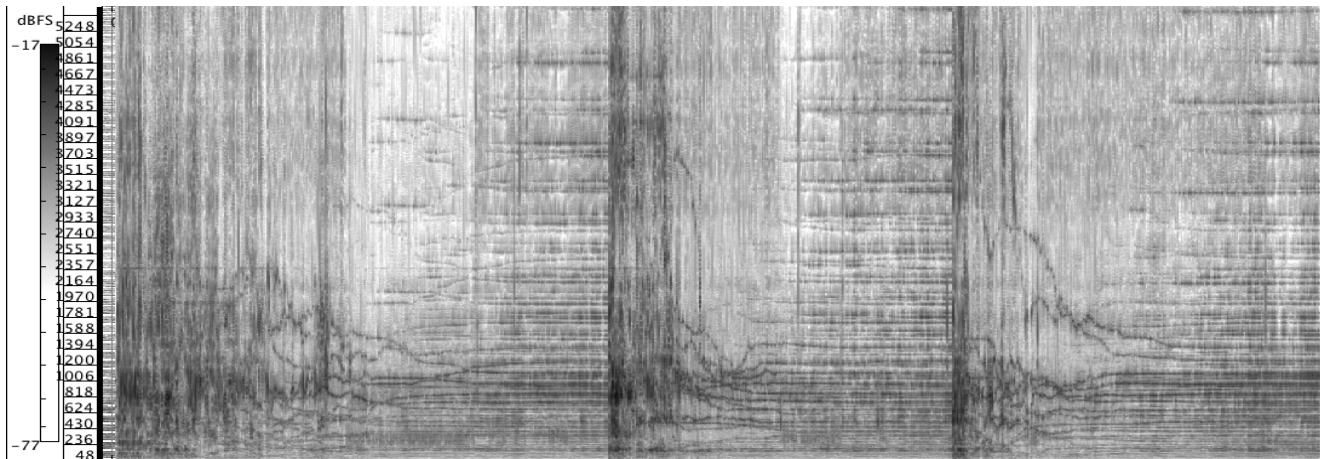


Figure 2: The converging optimization of three consecutive runs. The spectrogram shows the combined left and right channel. The effects from the mappings can be seen in the spectrum and are discussed in subsection 3.3.2.

in local minima. This exponential decay is reflected by the mutation strengths, which are encoded as a vector of σ components determining the variance of Gaussian mutation of the corresponding component of the solution.

We decided to map the mutation strengths to three different sonic parameters, namely filter-bandwidth, delay, and brightness. These three sound parameters were mapped with exponential mapping functions of different decay during optimization runs so that each effect would set in during different phases. You find a combined spectrogram for both stereo channels in figure 2, where the following effects can be studied:

The first effect, that gradually faded out during a run was the delay time for each of the 30 components. The effect was that the impressions of a big space with many different first reflections shrank to the acoustic impression of a small room.

The second effect was the decreasing bandwidth of the filters. Starting out with a big bandwidth the sum of all filters made sure that all spectral components of the typing sound passed through, and hence the augmentation was almost undistinguishable to the real typing sound. As the band width became smaller the filters started to exhibit a ring time that gave each component of the solution vector a noticeable characteristic in the stereo panorama.

The third effect was an increased brightness of the filter stacks, which was realized by lifting the level of the filters with the higher frequencies until all reached 0 dB. This added to the spectral contour lots of high frequency components. This effect was setting in during the last phase, when the optimization converged.

3.4. Audible effects of crucial optimization parameters

The purpose of monitoring optimizations is ultimately to tune them in real time if they do not converge. There are mainly two ways to influence the optimization of the ES algorithm during the run: control of the mutation strength, i.e. varying the σ during an optimization run. The effect of a changed mutation strength is clearly audible since it results in a changed bandwidth of the filters as well as a pronounce spatial impression of the sound through various delays and a noticeable difference in brightness.

Secondly, the selection pressure onto the population corresponding to the percentage of discarded individual solutions, can also be changed during a run. Tuning of this parameter becomes indirectly audible since it noticeably changes the algorithms dynamic, e.g. influencing the speed of movement towards the optimal solution. For some optimization cases the appropriately chosen selection pressure is crucial, if it is too low it would audibly result in a constantly changing, never converging sound pattern.

4. DISCUSSION

The combination of a one to many mapping with different decays ensured that the timbre of the augmentation was of noticeable difference when it matters, and with smaller difference at the beginning of the search process when the details of the algorithmic performance are of less interest. The sonic difference of different points in the parameter space were noticeable but subtle for points in the parameter space that were close to each other. However, given the fact that we used as test cases optimization problems of 30 dimensions, the result seems satisfying. One might not be able to remember the timbre of the convergence of the last run, but

a direct comparison of the end points in parameter space of several runs can help to judge if the points of convergence were the same.

4.1. Future Work

As next steps, we plan to extend our application so that after a set of optimization runs, the augmentation returns an auditory summary of all runs, thereby enabling an overview, if all converged to the same optimum. We further plan to extend the interactive monitoring to an integrated real-time control of optimizations. We also plan to develop dynamic scaling for the mapping from the points in search space. This is necessary to drive the whole setup into a direction, where it is applicable to potentially any continuous search problem addressed by ES. Further we will look at adaptive search algorithms; where meta-parameters will be an interesting target for real time interaction. Additionally, we plan to explore interaction possibilities for special optimization cases, where the space for optimal solutions is constrained.

4.2. Conclusion

Interactive sonification monitoring has shown interesting potential for monitoring ES algorithms by overcoming many shortcomings in existing displays. Interactive monitoring can easily turn from an unobtrusive indirect display with a PULL characteristic into an interactive direct display with a PUSH characteristic, where the user actively queries the state of convergence.

The real-time sonic representation through auditory augmentation allows to immediately monitor the success of convergence during optimization runs and offers an excellent way for the practitioner to become situated in a subtle parameter control feedback loop. The promising initial efforts encourage future research in order to adapt the approach for a more general applicability to evolutionary optimization algorithms.

5. ACKNOWLEDGMENT

We would like to thank Till Bovermann and Rene Tünnerman for inspiring exchange and discussions on auditory augmentation.

6. REFERENCES

- [1] F. Dombois, "Using audification in planetary seismology," in *Proceedings of the 7th International Conference on Auditory Display*, J. Hiipakka, N. Zacharov, and T. Takala, Eds., Laboratory of Acoustics and Audio Signal Processing and the Telecommunications Software and Multimedia Laboratory, Helsinki University of Technology. Espoo, Finland: ICAD, 2001, pp. 227–230. [Online]. Available: <http://www.icad.org/Proceedings/2001/Dombois2001.pdf>
- [2] P. Janat and E. Childs, "Marketbuzz: Sonification of real-time financial data," in *Proceedings of 10th Meeting of the International Conference on Auditory Display*, S. Barrass, Ed., International Community for Auditory Display (ICAD). Sydney, Australia: International Community for Auditory Display, 07 2004.
- [3] B. U. Rubin, "Audible information design in the new york city subway system: A case study," in *Proceedings of the International Conference on Auditory Display*, ICAD. British Computer Society, 1998. [Online]. Available: <http://www.icad.org/websiteV2.0/Conferences/ICAD98/icad98programme.html>
- [4] M. Ballora, B. Pennycook, P. C. Ivanov, L. Glass, and A. Goldberger, "Heart rate sonification: A new approach to medical diagnosis," *Leonardo*, vol. 37, no. 1, pp. 41 – 46, 2004.
- [5] S. V. Pereverzev, A. Loshak, S. Backhaus, J. C. Davis, and R. E. Packard, "Quantumoscillations between twoweaklycoupled reservoirs of superfluid 3he," *NATURE*, vol. 388, JULY 1997.
- [6] G. Baier, T. Hermann, and U. Stephani, "Multi-channel sonification of human eeg," in *Proceedings of the 13th International Conference on Auditory Display*, B. Martens, Ed., International Community for Auditory Display (ICAD). Montreal, Canada: ICAD, 06 2007, pp. 491–496.
- [7] T. Hermann, G. Baier, U. Stephani, and H. Ritter, "Vocal sonification of pathologic EEG features," in *Proceedings of the 12th International Conference on Auditory Display*, T. Stockman, Ed., International Community for Auditory Display (ICAD). London, UK: Department of Computer Science, Queen Mary, University of London UK, 06 2006, pp. 158–163.
- [8] A.S.Bregman, "Auditory scene analysis: The perceptual organization of sound," *TheMITPress*, September 1994.
- [9] H. G. Beyer and S. H. P., "Evolution strategies – a comprehensive introduction," *Natural Computing*, vol. 1, p. 352, 2002.
- [10] B. D. Fogel, *Artificial Intelligence through Simulated Evolution*. New York: Wiley, 1966.
- [11] J. H. Holland, *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, 1975.

- [12] I. Rechenberg, *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Stuttgart: FrommannHolzboog, 1973.
- [13] H. Schwefel, *Numerische Optimierung von Computer Modellen mittel der Evolutionsstrategie*. Basel: Birkhäuser, 1977.
- [14] T. Bartz-Beielstein, C. Lasarczyk, and M. Press, “Sequential parameter optimization,” in *Proceedings of the IEEE Congress on Evolutionary Computation CEC*, B. McKay *et al.*, Eds. IEEE Press, 2005, p. 773780.
- [15] O. Kramer, *Selfadaptive Heuristics for Evolutionary Computation*. Berlin: Springer, 2008.
- [16] S. Meyer-Nieberg and H. G. Beyer, *Parameter Setting in Evolutionary Algorithms editors*. Berlin: Springer, 2007, ch. Self Adaptation in Evolutionary Algorithms.
- [17] A. deCampo, C. Frauenberger, and R. Höldrich, “Designing a generalized sonification environment,” in *Proceedings of 10th Meeting of the International Conference on Auditory Display*. ICAD, 2004.
- [18] M. Wright, A. Freed, and A. Momeni, “Open sound control: State of the art 2003,” in *International Conference on New Interfaces for Musical Expression*, Montreal, 2003, pp. 153–159.
- [19] P. Vickers, “Chapter 18: Sonification for process monitoring,” in *The Sonification Handbook*, T. Hermann, A. Hunt, and J. Neuhoff, Eds. Berlin, Germany: Logos, Berlin, 2011.
- [20] T. Bovermann, R. Tünnermann, and T. Hermann, “Auditory augmentation,” *International Journal of Ambient Computing and Intelligence (IJACI)*, vol. 2, no. 2, pp. 27–41, 2010.
- [21] H. Schwefel, *Evolution and Optimum Seeking*. New York: Wiley Interscience Sixth Generation Computer Technology, 1995.