Using Exploratory Landscape Analysis to Visualize Single-objective Problems

Urban Škvorc Jožef Stefan Institute, Jožef Stefan International Postgraduate School Ljubljana, Slovenia urban.skvorc@ijs.si Tome Eftimov Jožef Stefan Institute Ljubljana, Slovenia tome.eftimov@ijs.si Peter Korošec Jožef Stefan Institute Ljubljana, Slovenia peter.korosec@ijs.si

ABSTRACT

Selecting the relevant algorithm for a given problem is a crucial first step for achieving good optimization algorithm performance. Exploratory Landscape Analysis can help with this problem by calculating landscape features that numerically describe individual problems.

To understand the problem space in single-objective numerical optimization, we recently presented a preliminary study of how Exploratory Landscape Analysis can be used to visualize different optimization benchmark problems, with the ultimate goal of visualizing problems that are similar to one another close together. In this paper, we examine how the selection of landscape features affects such a visualization, and show that proper preprocessing of landscape features is crucial for producing a good visualization. In particular, we show that only a subset of landscape features is invariant to simple transforms such as translation and scaling. Further, we examine how such an approach can be used to visually compare problems from different optimization benchmark sets.

CCS CONCEPTS

• Theory of computation → Continuous optimization; • Humancentered computing → Empirical studies in visualization; • Computing methodologies → Continuous space search;

KEYWORDS

Exploratory Landscape Analysis, Numerical Optimization, Visualization

ACM Reference Format:

Urban Škvorc, Tome Eftimov, and Peter Korošec. 2020. Using Exploratory Landscape Analysis to Visualize Single-objective Problems. In *Genetic and* Evolutionary Computation Conference Companion (GECCO '20 Companion), July 8–12, 2020, Cancún, Mexico. ACM, New York, NY, USA, 2 pages. https: //doi.org/10.1145/3377929.3397488

1 INTRODUCTION

In the field of optimization algorithms it is well known that the performance of any given optimization algorithm is highly dependent

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

GECCO '20, July 8-12, 2020, Cancun, Mexico

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-7127-8/20/07...\$15.00 https://doi.org/10.1145/3377929.3397488 on the problem this algorithm is solving. Because of this, determining methodologies to select the appropriate algorithm for a given task represents a critical focus of research in this field.

In order to determine which algorithm performs best at which problem, it is important to first understand the problems themselves.

Exploratory Landscape Analysis [6] is a method that allows us to describe an optimization problem using numeric descriptors called landscape features, with the idea that problems that are similar to each other will also have similar landscape features.

In our recently proposed study [8], we present a methodology that uses landscape features to visualize problems from two different single-objective optimization benchmarks, with the goal of visualizing similar problems close together. A particular focus is given to examining how specific landscape features are affected by common transformations that are applied to benchmark functions, such as translation and scaling.

Since different benchmarks apply different transformations, it is important that the visualization process is not affected by these transformations, as long as they do not have an effect on algorithm performance.

2 METHODOLOGY

To visualize the benchmark problems, we first collect samples from each function in a benchmark problem set using latin hypercube sampling with a sample size of 200 multiplied by the dimensionality of the problem. These samples are then used to calculate a set of landscape features for each of the problems using the open source library flacco [3]. To test our visualization, we use the problems from two single-objective numerical optimization benchmarks: the CEC Special Sessions and Competitions on Real-Parameter Single Objective optimization [4, 5], and the GECCO Black-Box Optimization Benchmark workshops [1].

Several preprocessing steps are then taken to improve the quality of the visualization. First, landscape features that are not invariant to transformations of scaling and shifting are removed. These two transformations are simple enough that they do not affect most state-of-the-art optimization algorithms (for example PSO and CMA) [2]), so any landscape features that are affected by them are not useful to us. This is done by taking a set of optimization problem samples, and then creating a shifted and scaled version of these samples for each problem in the benchmark set. Then, landscape features calculated on the original set are compared to the landscape features calculated on the transformed set using a paired Wilcoxon signed-rank test. The features that show a statistically significant difference between the original and the transformed function are removed from the set of landscape features and are not used for the visualization. After this step, all strongly correlated landscape features are also removed, and principal component analysis is used to reduce the dimensionality of the dataset in order to increase the quality of the visualization.

The t-Distributed Stochastic Neighbor Embedding [7] visualization method is then used to visualize the benchmark problems in two-dimensional space.

3 RESULTS

During the visualization process, we discovered that the transformations of scaling and translation have a large effect on some landscape features, which resulted in poor visualization. Removing the landscape features that were affected by these transformations greatly increased the quality of the visualization. The landscape feature groups that we found to be most useful for the visualization were *ela_distr, ela_meta, disp.ratio, pca.expl_var, cm_angle*, and *nbc.nb_fitness*.

The second experiment involved visualizing benchmark functions from two different years of the CEC benchmark competitions. These two sets contained several identical base functions that differed only in the transformations applied to them. In this experiment, we were able to visualize most of the identical functions close together.

The final experiment involved visualizing functions from two completely different benchmark sets: the BBOB workshops and CEC competitions. As in the previous experiment, these two benchmarks also shared some functions. Our method was able to visualize some of them close together, but failed in some cases. In order to examine why this happened, we directly plotted the samples of these functions and determined that there are significant differences in some shared functions between the two different benchmarks that could have been introduced because of different transformations. As a result some of the shared functions could not be visualized close together.

4 **DISCUSSION**

Our paper raises several points that we believe are of interest to the wider GECCO community. The most important is the fact that some landscape features are not invariant to simple transformations such as scaling and translation. Since most state-of-the-art optimization algorithms are invariant to such simple transformations, these non-invariant features should not be used if we want to compare optimization problems with the goal of determining which algorithms work best on which problems. Some optimization algorithms are also invariant to additional transformations such as rotation, so further work would be needed to analyze which landscape features are invariant to these transformations.

We also show that visualizing similar functions close together is much simpler when using problems from only a single benchmark set as opposed to two different benchmark sets. This indicates that there are either additional transformations that need to be taken into account (and landscape features that are not invariant to these transformations removed), or that there are other fundamental differences in the way different benchmarks use their benchmark problems, for example the domain of the problem that they are sampling from, or different problem parameters. However, if these changes have no effect on actual algorithm performance, they should also not have an influence on landscape features used for visualization.

We believe our research shows that Exploratory Landscape Analysis can be used to visually compare different benchmark problems. However, proper care must be taken to remove the influence of different benchmark methodologies (such as function transformations) on landscape features.

5 CONCLUSION

In the paper, we presented how Exploratory Landscape Analysis can be used for visualization of benchmark problems across different benchmark sets.

We have shown that only a subset of all landscape features offered by the popular library flacco can be used for this, as some features are not invariant to transformations that are commonly applied by benchmarks on their benchmark problems. However, further work is needed to determine which features are invariant to additional transformations, such as rotation, as well any other differences in benchmark design between different benchmark problem sets. Further work is also needed to determine how these transformations and other differences affect algorithm performance in practice.

ACKNOWLEDGMENTS

This work was supported by projects from the Slovenian Research Agency: research core funding No. P2-0098, project No. Z2-1867 and young researcher funding No. Pr-08987.

REFERENCES

- Nikolaus Hansen, Anne Auger, Dimo Brockhoff, Dejan Tusar, and Tea Tušar. 2016. COCO: Performance Assessment. (May 2016). https://hal.inria.fr/hal-01315318 ArXiv e-prints, arXiv:1605.03560.
- [2] Nikolaus Hansen, Raymond Ros, Nikolas Mauny, Marc Schoenauer, and Anne Auger. 2011. Impacts of invariance in search: When CMA-ES and PSO face illconditioned and non-separable problems. *Applied Soft Computing* 11, 8 (2011), 5755–5769.
- [3] Pascal Kerschke and Heike Trautmann. 2019. Comprehensive Feature-Based Landscape Analysis of Continuous and Constrained Optimization Problems Using the R-Package Flaccos. In Applications in Statistical Computing: From Music Data Analysis to Industrial Quality Improvement, Nadja Bauer, Katja Ickstadt, Karsten Lübke, Gero Szepannek, Heike Trautmann, and Maurizio Vichi (Eds.). Springer International Publishing, Cham, 93–123.
- [4] J. J. Liang, B. Y. Qu, and P. N. Suganthan. 2013. Problem definitions and evaluation criteria for the CEC 2014 special session and competition on single objective real-parameter numerical optimization. Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou China and Technical Report, Nanyang Technological University, Singapore (2013).
- [5] J. J. Liang, B. Y. Qu, P. N. Suganthan, and Q. Chen. 2014. Problem definitions and evaluation criteria for the CEC 2015 competition on learning-based real-parameter single objective optimization. Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou China and Technical Report, Nanyang Technological University, Singapore (2014).
- [6] Olaf Mersmann, Bernd Bischl, Heike Trautmann, Mike Preuss, Claus Weihs, and Günter Rudolph. 2011. Exploratory landscape analysis. In Proceedings of the 2011 Annual Conference on Genetic and Evolutionary Computation. ACM, 829–836.
- [7] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. Journal of machine learning research 9, Nov (2008), 2579–2605.
- [8] Urban Škvorc, Tome Eftimov, and Peter Korošec. 2020. Understanding the problem space in single-objective numerical optimization using exploratory landscape analysis. *Applied Soft Computing* 90 (2020), 106138. https://doi.org/10.1016/j.asoc. 2020.106138