Guerrero, J. L., Berlanga, A. & Molina, J. M. (2013). Multiobjective Local Search Techniques for Evolutionary Polygonal Approximation. In Omatu, S., et al. (eds.) *Distributed Computing and Artificial Intelligence: 10th International Conference*. (pp. 171-178). (Advances in Intelligent Systems and Computing; 217). Springer International Publishing.
DOI: http://dx.doi.org/10.1007/978-3-319-00551-5_21

Multiobjective Local Search Techniques for Evolutionary Polygonal Approximation

José L. Guerrero, Antonio Berlanga, and José M. Molina

University Carlos III of Madrid
Avda. Universidad Carlos III, 22, Colmenarejo, Spain
jguerrer@inf.uc3m.es, {aberlan,molina}@ia.uc3m.es

Abstract. Polygonal approximation is based on the division of a closed curve into a set of segments. This problem has been traditionally approached as a single-objective optimization issue where the representation error was minimized according to a set of restrictions and parameters. When these approaches try to be subsumed into more recent multi-objective ones, a number of issues arise. Current work successfully adapts two of these traditional approaches and introduces them as initialization procedures for a MOEA approach to polygonal approximation, being the results, both for initial and final fronts, analyzed according to their statistical significance over a set of traditional curves from the domain.

1 Introduction

Segmentation problems are based on the division of a given curve in a set of n segments (being each of these segments represented by a linear model, which points to another common naming convention for this process: piecewise linear representation, PLR) minimizing the representation error. Polygonal approximation techniques [10] are offline segmentation algorithms (since they require the whole curve they will be applied to) and can be divided into three different categories: sequential approaches, split and merge approaches and heuristic search approaches.

Sequential approaches are constructive methods based on a given local search over the current time series, trying to obtain, at each step, a new segment division (where the length of these segments is sequentially increased) which satisfies a certain criterion. Examples of the criteria used may be finding the longest possible segments ([12]). Split and merge approaches perform an initial segmentation over the given time series and afterwards start an iterative process to merge the initial segments until a certain criterion is met. According to their definition, these approaches have to deal with two different issues, the initial segmentation procedure and the merging criterion. An example of these processes is the bottom-up algorithm [7]

Heuristic search approaches are based on the development of heuristic methods in order to avoid the exhaustive search of the optimal *dominant points* for the given time series (which is a process with an exponential complexity). Different techniques may be used for this purpose, such as dynamic programming [11] or different metaheuristics, among them different solutions based on evolutionary algorithms [13]. The idea proposed by these works is to codify the time series as a chromosome with n genes, corresponding each of these genes to one of the points in the original data. If the gene value is a "1", it is considered a *dominant point*, and the algorithm tries to find the ideal codification of the chromosome according to a fitness function which evaluates the quality of the given codified segmentation in the chromosome.

Recently, the multiobjective nature of these processes is being explicitly approached from different perspectives [8, 5]. In [5] a multi-objective evolutionary algorithm [1] is proposed for the multi-objective solution of the segmentation issue, while in [4] a comparison between different possible initializations was carried, focusing on the different results between a random initialization aiming at the coverage of the obtained Pareto fronts versus the results from different local search techniques. One of the detailed issues is the single-objective nature of the traditional techniques used, which required different executions with different parameters in order to obtain different individuals from the front, also introducing issues regarding the configuration of these techniques to obtain such different individuals.

Current work will introduce a multi-objective explicit formulation from two traditional techniques for the polygonal approximation domain: bottom-up [7] and top-down algorithms[9]. This implementation will produce a whole Pareto front from a single execution without parametrization required from the user. These initializations will be later tested as initial populations for a MOEA approach, in order to determine whether they have successfully created better initial populations than the approach presented in [4] and how these initial populations translate into the final results of the algorithm.

The structure of this work is divided into three different sections: section two will present the techniques according to their traditional and multi-objective approach. After the new implementation has been presented, section three will present the results when these implementations are used to create the specified initial populations, analyzing the final results of the chosen algorithm. Finally the conclusions which can be extracted from the presented results will be presented in the final section, leading to the future lines of the work.

2 A Multiobjective Perspective to Local Search Polygonal Approximation Techniques

Two different issues can be stated regarding a polygonal approximation problem: $Min-\#$ and $Min-\epsilon$. $Min-\#$ is based on the optimization of the representation error for a previously set number of segments. $Min-\epsilon$, on the other hand, tries to find the minimum number of segments such that the final representation error does not exceed a previously established error ϵ. In [5] it was stated that, according to these two different perspectives, the segmentation issue is in fact a multi-objective problem, and also analyzed, according to different techniques available in the literature, how this nature had been faced. It was also shown that, given that some key dominant points are shared by different solutions with different resolutions, the solutions for $Min - \#$ and $Min - \epsilon$ problems can be closely related and share information between them. This multi-objective nature is faced with a Multi-objective evolutionary algorithm.

Local search algorithms may be introduced to enhance this approach, leading to several issues: the configuration to obtain the different individuals is hard to establish, and each of this individuals requires an independent execution of the local search algorithm, providing disappointing results [4]. This section will present alternative, parameter-free versions of two well known local search algorithms for polygonal approximation which provide a whole Pareto front of solutions: Top-Down and Bottom-up algorithms.

Top Down algorithm [9] is an offline process based on finding the best splitting point (understanding by this that measurement which divides the trajectory into the two segments with the lowest added errors) recursively, until all the resulting segments have an error value bellow a user defined boundary. The Top Down algorithm is applied in a wide variety of domains and fields, being also known by different names[2].

The multi-objective version of the Top Down algorithm suppresses the two issues available in the traditional implementation: the recursive calls (which may prevent the application of the algorithm to figures with a large number of points) and the user configuration (which introduces the issues previously described in the obtaining of a whole Pareto front). At each step, the best splitting point is located (the one which provides with the smallest representation error), a new individual is generated adding that new dominant point and the costs of the possible segments are updated (implying the recomputation of the costs of the segments from the dominant point immediately to the left of the new splitting point and those from the splitting point to the dominant one immediately to its right). Therefore, no recursive calls are included, and each split point choice has a global view of the representation error (as opposed to the partial one available in the traditional implementation). Figure 1 represents the multi-objective version implementation of this algorithm.

Bottom up algorithm[7] is an offline process complementary to Top Down, where the time series is initially divided into every possible segment (composed of two measurements) and finds the best possible segment fusion afterwards (understanding by this the fusion which obtains the segment with the lowest error) until any possible fusion obtains a segment having an error above a user defined boundary. The bottom up algorithm, as well, has

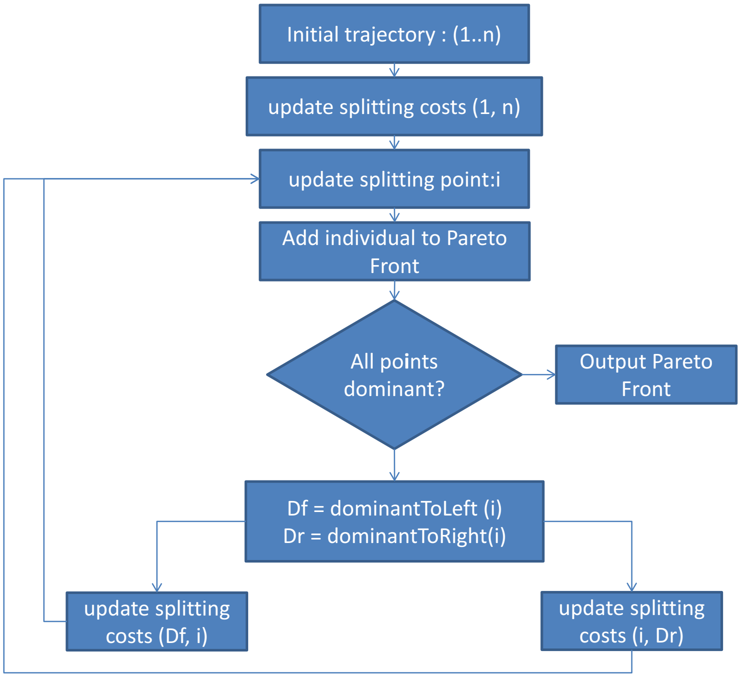


Fig. 1 Top Down algorithm multi-objective implementation

spread to different fields and research areas using different names, such as the computer graphics domain and decimation methods[6].

The multi-objective version of bottom-up algorithm removes the user-defined boundaries for the algorithm termination, being this ending triggered once no further merging can be performed. Figure 2 presents the multi-objective version. It must be noted that each update here triggers only one segment update, while every new splitting point in the top down algorithm triggered the recomputation of all the possible new splitting points for the two new segments created in the representation. Since each of these steps are, in fact, mutations over the chromosome guided by a specific heuristic, the principles for an efficient implementation established in [3] can be applied for the computation of the fitness values of each of the produced individuals.

3 Experimental Validation: Initialization for MOEA Polygonal Approximation

The experimental validation proposed will include the two detailed multi-objective local search procedures to create the initial populations for a Multi-objective evolutionary approach to polygonal approximation. This algorithm

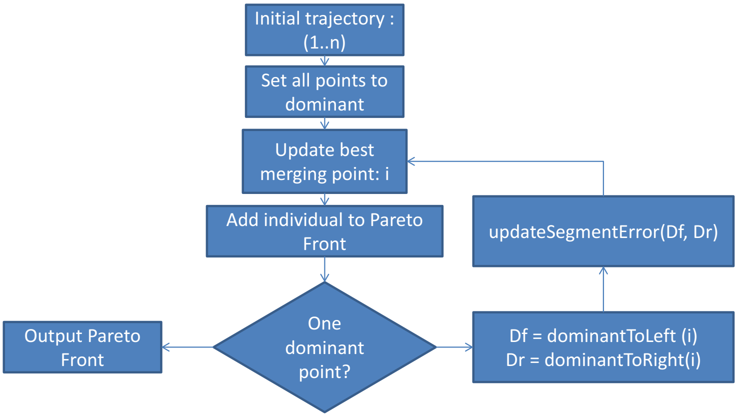
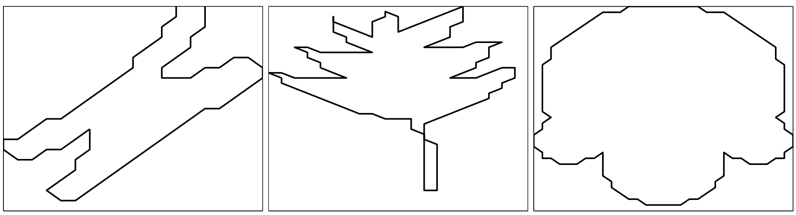


Fig. 2 Bottom up algorithm multi-objective implementation

is based on the SPEA2 [14] MOEA, according to the configuration presented in [5]. The default initialization process creates a uniform Pareto Front in terms of coverage of the objectives, as presented in [4]. This section will cover the comparison between the initial and final populations of the two techniques presented and the suggested initialization process. The dataset used is composed of three traditional curves, usually named chromosome, leaf and semicicle. Their definition, according to ther freeman chain-code representation, can be found in [5]. Figure 3 represents these figures.



(a) Chromosome curve (b) Leaf curve (c) Semicircle curve

Fig. 3 Curves included in the data set

It is interesting to notice the complementary nature of the two multi-objective techniques presented, since once applied its heuristic with a value of 1 dominant point and applies successive splitting over the figure (Top-Down) and the other begins with a solution with all of its points considered dominant and applies successive merging (Bottom-up). Since the solutions tend to degrade with the successive application of the heuristic, each of them will be more successful at their initial individuals. This can be seen in figure

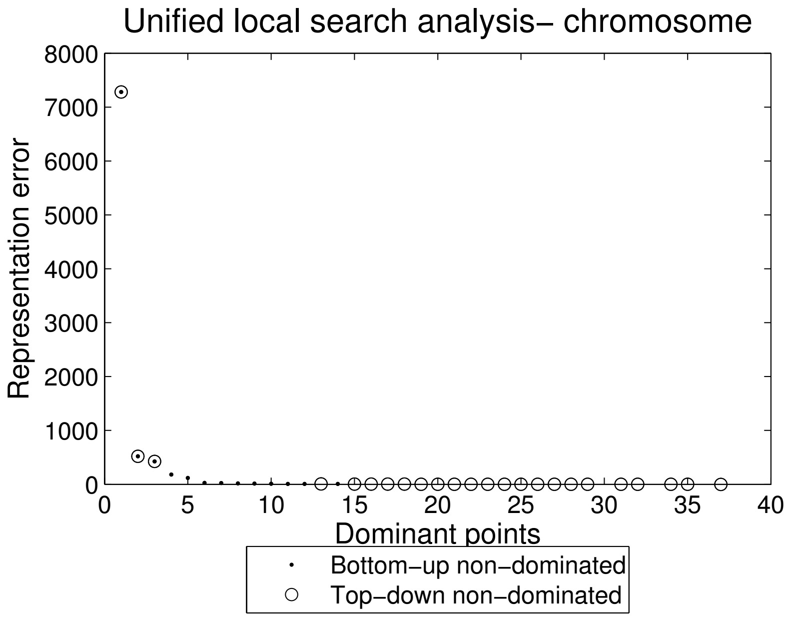


Fig. 4 Analysis of unified local search front non-dominated individuals

4 which shown the non-dominated individuals of a Pareto front composed of bottom-up and top-down initial fronts for the chromosome curve (figure 3a).

The results for the four techniques, including their mean and median values for the hypervolume of the obtained Pareto fronts are included in tables 1 (initial fronts values) and 2(final fronts values). Also, a best technique column has been added. This value is calculated according to a Wilcoxon test with a 95% confidence performed over 30 different executions, since the values do not follow a normal distribution. If one technique is superior to the remaining ones, its name is included, otherwise the '-' value is included.

The results show that, on the one hand, the multi-objective search approach is able to provide better initial populations in terms of hypervolume. When these populations are used by the underlying MOEA algorithm, different cases appear. For easy problems, such as Chromosome, the uniform initialization provides better final results. This happens due to the focus which the local search initialization introduces according to its underlying heuristic. Even though the initial results are clearly improved, the final ones are too guided by the initial heuristic. For harder problems, the initialization provides the algorithm with an important enough advantage such that the final populations are either not statistically significant (leaf) or significantly better (chromosome). These results seem to point to a combination of both techniques to provide initial populations that, while benefiting from the

Table 1 Initial populations comparison

Figure	Bottom-up Mean	Bottom-up Median	Top-down Mean	Top-down Median	Local search Mean	Local search Median	Uniform Mean	Uniform Median	Best
Chrom.	0,98647	0,98647	0,98646	0,98646	0,98651	0,98651	0,98436	0,98427	L.S.
Leaf	0,99355	0,99355	0,99322	0,99322	0,99365	0,99365	0,99271	0,99281	L.S.
Semi.	0,99157	0,99157	0,99183	0,99183	0,99218	0,99218	0,99101	0,99111	L.S.

Table 2 Final populations comparison

Figure	Bottom-up Mean	Bottom-up Median	Top-down Mean	Top-down Median	Local search Mean	Local search Median	Uniform Mean	Uniform Median	Best
Chrom.	0.98665	0.98664	0.98667	0.98671	0.98665	0.98667	0.98671	0.98672	Unif.
Leaf	0.99376	0.99376	0.99374	0.99376	0.99376	0.99376	0.99377	0.99378	-
Semi.	0.99206	0.99219	0.99213	0.99217	0.99219	0.99219	0.99213	0.99217	L.S.

enhanced initial populations of local search techniques, can be not hampered by the heuristic focus.

4 Conclusions

Local search techniques have been the focus of polygonal approximation, developing different techniques based on specific heuristics for this issue. However, the new multi-objective approaches require modifications over these techniques in order to efficiently obtained the required Pareto Fronts. This work has modified Bottom-up and Top-down techniques to provide a multi-objective approach with the required characteristics presented. The results have been tested for the initialization of a MOEA algorithm. These results show that the multi-objective techniques are successful in providing statistically better initial populations, however the final results may be too focused on the heuristic used in these techniques, which in some cases hamper their quality. Future lines imply the research of the combination which may be performed over local-search and uniform initialization in order to provide initial populations taking advantage of local search improved initial populations without their excessive focus on their underlying heuristic.

Acknowledgement. This work was supported in part by Projects MEyC TEC2012-37832-C02-01, MEyC TEC2011-28626-C02-02 and CAM CONTEXTS (S2009/TIC-1485).

References

1. Coello, C., Lamont, G., Van Veldhuizen, D.: Evolutionary algorithms for solving multi-objective problems. Springer-Verlag New York Inc. (2007)
2. Duda, R.O., Hart, P.E.: Pattern Classification and Scene Analysis. Wiley (1973)
3. Guerrero, J., Berlanga, A., Molina, J.: Fitness-aware operators for evolutionary polygonal approximation. In: Applied Computing, pp. 283–290. IADIS (2012)
4. Guerrero, J.L., Berlanga, A., Molina, J.M.: Initialization procedures for multiobjective evolutionary approaches to the segmentation issue. In: Corchado, E., Snášel, V., Abraham, A., Woźniak, M., Graña, M., Cho, S.-B. (eds.) HAIS 2012, Part III. LNCS, vol. 7208, pp. 452–463. Springer, Heidelberg (2012)
5. Guerrero, J., Berlanga, A., Molina, J.: A multi-objective approach for the segmentation issue. Engineering Optimization 44(3), 267–287 (2012)
6. Heckbert, P., Garland, M.: Survey of polygonal surface simplification algorithms. In: Multiresolution Surface Modeling Course. ACM Siggraph Course notes (1997)
7. Keogh, E., Chu, S., Hart, D., Pazzani, M.: Segmenting time series: A survey and novel approach. Data Mining in time series databases, 1–21 (2004)
8. Kolesnikov, A., Franti, P., Wu, X.: Multiresolution polygonal approximation of digital curves. In: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 2, pp. 855–858. IEEE (2004)
9. Ramer, U.: An iterative procedure for the polygonal approximation of plane curves. Computer Graphics and Image Processing 1, 244–256 (1972)
10. Sarfraz, M.: Linear capture of digital curves. In: Interactive Curve Modeling, pp. 241–265. Springer, London (2008)
11. Sato, Y.: Piecewise linear approximation of plane curves by perimeter optimization. Pattern Recognition 25(12), 1535–1543 (1992)
12. Sklansky, J., Gonzalez, V.: Fast polygonal approximation of digitized curves. Pattern Recognition 12(5), 327–331 (1980)
13. Yin, P.: Genetic algorithms for polygonal approximation of digital curves. International Journal of Pattern Recognition and Artificial Intelligence 13(7), 1061–1082 (1999)
14. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the Strength Pareto Evolutionary Algorithm. In: EUROGEN 2001, pp. 95–100. International Center for Numerical Methods in Engineering (CIMNE), Athens (2001)