Bioinformatics, 34(3), 2018, 538–540
doi: 10.1093/bioinformatics/btx621
Advance Access Publication Date: 27 September 2017
Applications Note

OXFORD

Bioimage informatics

DiversePathsJ: diverse shortest paths for bioimage analysis

Virginie Uhlmann[1], Carsten Haubold[2], Fred A. Hamprecht[2,*] and Michael Unser[1,*]

[1]École polytechnique fédérale de Lausanne (EPFL), Biomedical Imaging Group, 1015 Lausanne, Switzerland and [2]Heidelberg University, Interdisziplinären Zentrums für Wissenschaftliches Rechnen/Heidelberg Collaboratory for Image Processing (IWR/HCI), 69120 Heidelberg, Germany

*To whom correspondence should be addressed.

Associate Editor: Robert Murphy

Received on July 12, 2017; revised on September 19, 2017; editorial decision on September 24, 2017; accepted on September 26, 2017

Abstract

Motivation: We introduce a formulation for the general task of finding diverse shortest paths between two end-points. Our approach is not linked to a specific biological problem and can be applied to a large variety of images thanks to its generic implementation as a user-friendly ImageJ/Fiji plugin. It relies on the introduction of additional layers in a Viterbi path graph, which requires slight modifications to the standard Viterbi algorithm rules. This layered graph construction allows for the specification of various constraints imposing diversity between solutions.
Results: The software allows obtaining a collection of diverse shortest paths under some user-defined constraints through a convenient and user-friendly interface. It can be used alone or be integrated into larger image analysis pipelines.
Availability and implementation: http://bigwww.epfl.ch/algorithms/diversepathsj
Contact: michael.unser@epfl.ch or fred.hamprecht@iwr.uni-heidelberg.de
Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

A large number of problems in image analysis and computer vision can be formulated as search for the *shortest path*. Given an edge-weighted graph, the optimal shortest path between two nodes can be found efficiently using dynamic programming (Bellman, 1952). The Viterbi algorithm is a popular example of the latter. A solution is, however, always only optimal with respect to its associated cost function or energy. In practice, due to discrepancies between the model associated with an objective function and reality, it is common for the *optimal* solution to differ from what would be the *best* one from a human perspective. For this reason, obtaining a collection of good solutions instead of a single one is often desirable.

The design of algorithms for obtaining the M most probable solutions to combinatorial problems has been extensively studied in the context of probabilistic random field models and is known as the M-best maximum a posteriori (MAP) problem (Batra, 2012; Lawler, 1972). Further works aimed at introducing supplementary constraints to ensure that the M-best solutions are sufficiently diverse (Batra *et al.*, 2012; Kirillov *et al.*, 2015). In Straehle *et al.* (2013), these very general formulations were adapted for bioimage segmentation in the context of exploring various watershed solutions by searching for the K-smallest spanning trees. Most similar to our approach is the one from Fujita *et al.* (2003) in the context of robot motion planning, which could also be viewed as a two-layers-graph construction.

We make two contributions: first, we propose a more general formulation for the diversity constraint and second, we provide the first freely available, readily usable and general purpose open source implementation of a diverse shortest path tool for bioimage analysis.

2 Materials and methods

The search for a diverse collection of M-diverse best solutions using a layered-graph construction is implemented efficiently by modifying the

© The Author 2017. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/34/3/538/4259193 by Universite and EPFL Lausanne user on 18 January 2019

rules of the standard Viterbi algorithm. The mathematical formulation of our approach is inspired from our more general work (Haubold *et al.*, 2017) and is provided in the Supplementary Material. Here, we give an intuitive description of the algorithm.

To find the shortest path, the standard Viterbi algorithm is performed from a source point σ to a target point τ by computing the cost to all intermediate nodes. We refer to this collection of costs as the *forward* derived cost map (Fig. 1a). The shortest path is found by backtracking the minimal cost from τ to σ (Fig. 1b). For the second shortest path, another pass of Viterbi algorithm is first performed, this time from the target τ to the source σ, to obtain the *backward* derived cost map (Fig. 1c). The forward and backward derived cost map are then summed (Fig. 1d) and all nodes laying within an *exclusion corridor* around the shortest path are assigned an infinitely large cost (represented as white area in Fig. 1e). The exclusion corridor aims at enforcing that the second shortest path differs from the first one. The node corresponding to the new minimal cost in the summed map, denoted as $(\nu^*, x^*_{\nu^*})$, is identified. Finally, the second diverse shortest path is obtained by joining the two sub-paths obtained by backtracking from $(\nu^*, x^*_{\nu^*})$ to σ and from $(\nu^*, x^*_{\nu^*})$ to τ. This amounts to performing the Viterbi algorithm on a two-layers graph constructed by copying the original graph and introducing layer-jump-edges (Fig. 1f).

The next $M-2$ diverse best solutions can be computed efficiently by updating the exclusion corridor to be the union of the exclusion corridors built from previously found solutions, and searching for the next minimum in the remaining available nodes of the summed cost map. The M shortest paths are thus computed at the cost of two standard Viterbi algorithm passes plus M backtrackings. A possible refinement is to enforce in addition that the new solution accumulates a given amount of diversity. Formally, this requires the introduction of a supplementary input cost map on the lower graph layer encoding the discrepancy between the new solution and previous ones, refereed to as the *diversity* map. It can be combined with the exclusion corridor to enhance diversity in the set of shortest paths. Layer-jump-edges are then allowed only when the integrated diversity over the path exceeds a threshold.

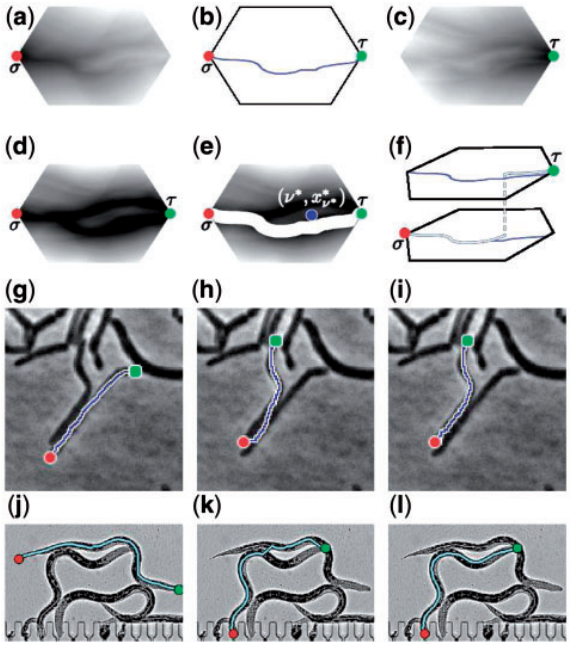

Fig. 1. *Algorithm for finding the second best path.* (**a**) Forward derived cost map (from source node σ), (**b**) shortest path, (**c**) backward derived cost map (from target node τ), (**d**) summed cost map, (**e**) exclusion corridor mask and node $(\nu^*, x^*_{\nu^*})$ with minimal cost and (**f**) resulting second shortest path obtained by backtracking paths from $(\nu^*, x^*_{\nu^*})$ to σ and τ. The underlying input image, source and target locations are those shown in (**k**). *Application examples in images of mycobacteria (phase-contrast) and nematodes (bright-field).* For some objects, the shortest path corresponds to the desired one [(**g**) and (**j**)], but for others in the same image [(**h**) and (**k**)], only the Mth shortest is correct [(**i**) and (**l**)]

Our approach is implemented as a Java-based plug-in for ImageJ and Fiji. The plug-in is designed to be generic and readily usable. A full description of the GUI and all available settings is provided in the Supplementary Material. The required input consists in two end-points provided by the user or by custom automated pre-processing steps. The collection of resulting (diverse) paths can then either be swiftly browsed and displayed on the image using the arrow keyboard keys, or exported in generic formats for further processing.

3 Results

Bioimages often feature several instances of the same object in a single image (e.g. cells). Due to small differences between the appearance and neighborhood of these individual objects, it is practically impossible to find a cost function whose minimum always corresponds to the desired output. Exploring several shortest paths, however, allows to analyze multiple instances of the same objects exhibiting slight variations without the need for fine-tuning of the cost function.

We illustrate the usefulness of searching for several shortest paths in bioimage analysis with identification of the medial axis of mycobacteria (phase contrast microscopy images) and of *C.elegans* nematodes (brightfield microscopy images). The original images, their characteristic, experimental protocols and plugin settings can be found in the Supplementary Material. We use our approach with the same parameters to analyze two similar objects in the same image. For one of the objects (Fig. 1g and j), the shortest path is the desired medial axis. Conversely, for the second one, the shortest path does not correspond to a correct medial axis (Fig. 1h and k). The desired solution can, however, be found in the set of additional diverse shortest path (Fig. 1i and l), the exploration of which is made easy through the interface of our ImageJ/Fiji plugin. A full description of the software is provided in the Supplementary Material.

Medial axis outline serves as the first step for quantitative estimation of shape features such as length and bending. The proposed approach is however not restricted to this use-case and covers every problem in which a path is to be searched between two end-points. Our implementation only requires the specification of the source and target locations as ImageJ points and of the input cost map specifying the cost associated to each pixel in the image, in addition to data-independent parameters (number of desired shortest paths and minimum amount of diversity between paths). Ridge detectors and algorithms estimating probability distributions for the identification of principal curves such as subspace constrained mean shift (Ozertem *et al.*, 2011) are good examples of relevant input cost maps. DiversePathsJ can be used alone in an interactive way or integrated as a module in larger automated image analysis pipelines.

Funding

This work was supported by the Swiss National Science Foundation under Grant 200020_162343/1 (VU and MU); and by DFG SFB1129, DFG HA4364/9-1 and HGS MathComp (CH and FAH).

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/34/3/538/4259193 by Universite and EPFL Lausanne user on 18 January 2019

Conflict of Interest: none declared.

References

Batra,D. (2012) An efficient message-passing algorithm for the M-best MAP problem. In: *Proc. 28th Conf. Uncertain. Artif. Intell. (UAI'12)*, Catalina Island, USA, pp. 121–130.

Batra,D. *et al.* (2012) Diverse M-best solutions in markov random fields. In: *Proc. 12th Eur. Conf. Comput. Vis. (ECCV'12)*, Florence, Italy, pp. 1–16.

Bellman,R. (1952) On the theory of dynamic programming. *Proc. Natl. Acad. Sci. USA*, **38**, 716–719.

Fujita,Y. *et al.* (2003) Dual Dijkstra search for paths with different topologies. In: *Proc. IEEE Int. Conf. Robotic. Autom. (ICRA'03)*, vol. **3**, Taipei, Taiwan, pp. 3359–3364.

Haubold,C. *et al.* (2017) Diverse *M*-best solutions by dynamic programming. In: *Proc. 39th Ger. Conf. Pattern Recognit. (GCPR'17)*, Basel, Switzerland.

Kirillov,A. *et al.* (2015) Inferring M-best diverse labelings in a single one. In: *Proc. IEEE Int. Conf. Comput. Vis. (ICCV'15)*, Santiago, Chile, pp. 1814–1822.

Lawler,E. (1972) A procedure for computing the k best solutions to discrete optimization problems and its application to the shortest path problem. *Manag. Sci.*, **18**, 401–405.

Ozertem,U. *et al.* (2011) Locally defined principal curves and surfaces. *J. Mach. Learn. Res.*, **12**, 1249–1286.

Straehle,C. *et al.* (2013) K-smallest spanning tree segmentations. In: *Proc. 35th Ger. Conf. Pattern Recognit. (GCPR'13)*, Saarbrücken, Germany, pp. 375–384.

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/34/3/538/4259193 by Universite and EPFL Lausanne user on 18 January 2019