CVPR
#5813

CVPR
#5813

CVPR 2020 Submission #5813. CONFIDENTIAL REVIEW COPY. DO NOT DISTRIBUTE.

Rebuttal: GASP, a generalized framework for agglomerative clustering of signed graphs and its application to Instance Segmentation

We thank the three reviewers for their helpful comments and for pointing out that our manuscript is very clear and well written (R1, R2, R3), is theoretically sound and provides a novel level of abstraction (R1) and is highly valuable in practice by achieving competitive accuracies on two publicly available benchmarks (R1, R3).

Novelty and main contributions (R2, R3) – Our main contribution is GASP, a framework that generalizes several existing and novel algorithms for agglomerative clustering of signed graphs, providing a powerful tool to study these methods. R2 states that *"exactly the same algorithms as GASP have been proposed"*, referring to GF [54] and GAEC [40]. Indeed, GF and GAEC are special cases of GASP, as we point out in Table 1 and lines 369-371. However, they are only two out of 14 algorithms described in our framework. In this paper we prove, for the first time, that Mutex Watershed [90] is another specialization of GASP – an insight which we claim is nontrivial – but all other algorithms are new at all, or at least in the context of signed graph partitioning. Indeed, one of the new algorithms, which on signed graphs has not been studied before, turns out to be the best among all variants we explored.

R2 also states that *"extension to include other linkage strategies is trivial"*. We agree wholeheartedly, and indeed consider this a great strength of the proposed framework, which allows an easy classification and comparison of agglomerative algorithms for signed graph clustering.

R3 points out that the use of signed graph clustering for instance segmentation is not novel and has been applied both in the context of natural images and EM connectomics. We agree, and cite pertinent prior work [42, 2, 7, 94]. Rather, we introduce a new framework that generalizes several existing methods and yields variants that have not been studied before. We believe that our contribution generates new insights into the utility of agglomerative clustering methods for Computer Vision and show their competitive performance when combined with predictions from neural networks, despite the fact that they are simple and efficient compared to other methods for signed graph partitioning.

Comparison to multicut solvers (R2) – In lines 75-80 we write that *"Agglomerative clustering algorithms for signed graphs have clear advantages: [...]. Despite the fact that a variety of these algorithms exist, no overarching study so far has been conducted to compare their robustness [...]"*. To the best of our knowledge, this claim pertaining to agglomerative algorithms is correct, and R2 does not offer references to the contrary. As R2 correctly points out, [54] contains an extensive evaluation of local search algorithms as multicut approximations, including the agglomerative methods GAEC and GF. However, [54] does not make a connection to agglomerative methods with different linkage criteria studied here. Furthermore, R2 mentions the comparison of multicut solvers in [Beier et al., Cut, Glue & Cut: A Fast, Approximate Solver for Multicut Partitioning]. We are aware of this work, but did not consider it sufficiently relevant in the context of our contribution, because it does not contain any comparison with agglomerative methods. Still, in the revised version we will cite it following R2.

Experiments on larger graphs (R2) – R2 points out that we should conduct experiments on a broader set of problems and use graphs of larger size, referring to the experiments in [54]. We note that our experiments cover two important applications of instance segmentation from very different images (Cityscapes vs electron microscopy images) as well as clustering on random graphs, where we compare favorably to spectral methods. Regarding the problem size, note that the CREMI problem used for evaluation (Table 2) is solved on the pixel grid graph and contains $7 \cdot 10^7$ nodes and $6 \cdot 10^8$ edges, which is two orders of magnitude larger than the largest instance segmentation problem in [54] consisting of 10^5 nodes and 10^7 edges. In the revised version of the paper, we will add the specifications of the problem size on CREMI.

Relevance of the results on CREMI (R2) – In the revised version, we will better define the CREMI score, which is given by the geometric mean of adapted Rand Error and Variation of Information, two commonly used metrics to evaluate clusterings. Given the small gap of 0.02 to the leading submission and the much bigger gap to other published methods, e.g. more than 0.3 gap for [95], it is fair to claim that we are competitive with state-of-the-art. In the revised version of the paper, we will comment on our standing in the CREMI leader-board.

Clarifications about cityscapes experiments (R2) – R2 asks about the difference between the agglomeration method used in GMIS [63] and GASP. GMIS uses average linkage agglomerative clustering on an *unsigned graph*, whereas GASP operates on a *signed graph*. In addition, as we point out in lines 769-771, the agglomeration procedure in GMIS is applied in multiple stages, which creates the need for further hyper-parameters and makes it more complex than our algorithm, which achieves comparable (slightly better) accuracy.

We have fine tuned the GMIS affinity predictions using the Dice Loss, which is a common approach to improve results for sparse spatial targets like affinities, see [90].