Review on “Global Data Association for Multi-Object Tracking Using Network Flows”

Authors: L. Zhang, Y. Li, R. Nevatia

# Short Summary

This paper formulates an approach for global data association that outperforms previous methods on object tracking in long sequences. Specifically, data association is presented as a maximum-a-posteriori (MAP) problem mapped into a cost-flow network. Associated non-overlapping trajectories are modelled as disjoint flow paths where observation likelihood and transition probabilities are flow costs. Thus, the optimal trajectory can be found by a minimum-cost flow algorithm (Algorithm 1). The paper also discusses the inclusion of occlusion nodes and constraints as an Explicit Occlusion Model (EOM) which mitigates the impact of long-term occlusion. This new network can be solved using an iterative approach built upon the original minimum-cost flow algorithm (Algorithm 2). Uniquely, since the method contains trajectory initialization, termination and object occlusions, these parameters can be inferred from the solution.

The final two algorithms are evaluated on the CAVIAR and ETHMS video datasets which suffer from high occlusion and poor image contrast. The second algorithm improves detection performance at the expensive of a higher false positive rate. Both methods demonstrate performance comparable to a well-established baseline which at the time of publishing scored highest on the CAVIAR dataset. Additionally, the proposed algorithms reduce false positives. Despite a polynomial theoretical complexity, the model complexity seems to grow linearly in practice.

# Main Contributions

1. Proposes a novel data association framework for multiple object tracking that optimizes association globally using the entire sequence of observations
2. Proposes the inclusion of an Explicit Occlusion Model (EOM) to deal with long-term occlusions
3. Proposes an iterative solution methodology using a modified min-cost flow algorithm

# High-Level Evaluation of Paper

Of all the papers I’ve read thus far, I struggled the most with following this paper. Although a high-level intuition for the methodology is provided, the supporting mathematics is difficult to follow and requires familiarity with dynamic programming or optimization. Additionally, I do not have much familiarity with this task (object tracking) which further complicated the reading. I feel that the authors should have opted to extend the length of the paper (7 pages) to provide concise explanations for what intermediate mathematical steps they are performing and why.

# Discussion on Evaluation Methodology

The initial evaluation where the proposed Algorithms are compared to Wu *et al.* [2 in paper] on the CAVIAR dataset provides insight into trade-offs. Evaluated on five metrics, one can identify readily in what ways the first proposed algorithm builds on the baseline and then the second algorithm builds on the baseline. Qualitative results are also provided in a figure to give a concrete example of the benefits in practice. The more detailed second table compares detection performance and false alarms per frame to similarly contrast the baselines with the proposals and is effective in conveying the results.

The authors strangely opted to exclude a table summarizing the complexity of both models and additionally do not compare this important aspect to any historical baselines. It’s strange that they bring this up as a benefit yet do not develop the idea.