| | Method | Param (M) ↓ | Sintel (val) EPE ↓ | |
|---|---|---|---|---|
| | | | Clean | Final |
| Grid Feature | RAFT | 5.26 | 1.65 | 3.04 |
| | + Dense Convs | + 3.34 | 1.63 | 2.98 |
| | + Dilated Convs | + 0.85 | 1.66 | 3.02 |
| | + SuperGlue | + 2.99 | - | - |
| Graph Space | GCU (Base) | 5.44 | 1.76 | 3.14 |
| | + G-SuperGlue | + 1.82 | 1.63 | 3.01 |
| | + **AGR (ours)** | + 0.12 | **1.50** | **2.88** |

Table 2: Quantitative comparisons with related methods (refer to Sec.  for more details). We set RAFT (**?**) as the baseline model for the method with regular gird space enhancement. The methods in each part are plugged into the same baseline and trained on C + T ($180k$) for fair comparison.

| | Settings | FLOPs (G) | Param (M) | Sintel (val) EPE | |
|---|---|---|---|---|---|
| | | | | Clean | Final |
| Graph Reasoning | Base Graph | 381.06 | 5.44 | 1.76 | 3.14 |
| | + SGR (no GA) | + 16.95 | + 0.07 | 1.65 | 2.97 |
| | <u>+ AGR</u> | + 17.09 | + 0.12 | 1.50 | 2.88 |
| Node numbers | $K = 32$ | + 3.09 | + 0.04 | 1.61 | 3.02 |
| | $K = 64$ | + 5.88 | + 0.07 | 1.56 | 2.95 |
| | <u>$K = 128$</u> | + 17.09 | + 0.12 | 1.50 | 2.88 |
| | $K = 256$ | + 61.84 | + 0.2 | 1.49 | 2.90 |
| Attentive readout | <u>On</u> | + 17.09 | + 0.12 | 1.50 | 2.88 |
| | Off | + 16.82 | + 0.10 | 1.55 | 2.94 |

Table 3: Ablation analysis for different settings of our AGFlow. "SGR" indicates separated graph reasoning for context and motion nodes (*i.e.*, without graph adapter), and "AGR" denotes overall adaptive graph reasoning. All methods are trained on C + T ($180k$) for fair comparison. <u>Underline</u> indicates the default settings in our model.

in a performance drop on both passes of Sintel. We regard it as the base graph model, as in line 5 of Tab. 2. In contrast, we carefully project feature maps into context and motion nodes, and further propose an adaptive graph reasoning approach to perform the *task-specific* hybrid reasoning, allowing our model to significantly reduce the average end-point error by around $14.8\%$.

**Effectiveness of Adaptive Graph Reasoning.** In Tab. 3, we empirically analyze the computational cost and corresponding performance gain of the core component in the proposed AGFlow. Using separated context and motion reasoning based on graph model boosts the performance by $5.4\% \sim 6.3\%$. Besides, we further incorporate the proposed graph adapter into the context-to-motion interaction, and then yield the proposed adaptive graph reasoning (AGR) module, which brings about $9\%$ additional performance gain with only $0.14$ G FLOPs and $0.05$ M parameters for extra computational overhead.

**Ablation for Node Numbers.** We empirically show the influences of node numbers $K$ in our graph model. As shown in Tab. 3, when more nodes are used ($32 \rightarrow 64 \rightarrow 128$), the average end-point error are gradually decreased from $1.61$ on Sintel clean pass and $3.02$ on final pass to $1.50$ and $2.88$, respectively. However, if furthermore nodes are involved ($128 \rightarrow 256$), the flow accuracy almost remains the same and the computational overhead is largely increased by $2.6$ times. This is because some redundant feature representations are generated with nodes, which brings no benefit to flow estimation. Therefore, we set $K = 128$ to ensure a good balance between efficiency and performance.

| Method | Param (M) ↓ | Time (ms) ↓ |
|---|---|---|
| RAFT | 5.26 (-) | 86.9 (-) |
| GMA | 5.8 (+ 0.54) | 113.8 (+ 26.9) |
| **AGFlow** | 5.56 (+ 0.30) | 90.7 (+ 3.8) |

Table 4: Computational comparisons with state-of-the-arts on a single Geforce RTX 2080Ti GPU.

**Ablation for Attentive Readout.** We also test the influence of attentive readout compared with regular readouts (**?**) in Tab. 3. As can be seen, incorporating it into our model brings about $3\%$ in performance gain and only requires negligible computation cost and parameters, demonstrating the cost-effective property of this component.

**Runtime Comparison.** We provide the parameters and runtime of state-of-the-art methods in Tab. 4. Compared with GMA (**?**), our AGFlow can achieve competitive performance while reducing $0.24$ M parameters. Besides, the inference speed is boosted by $20.3\%$ ($113.8 \rightarrow 90.7$). The comparisons clearly demonstrate the effectiveness of our AGFlow.

## Conclusion

In this paper, we present a novel graph-based approach termed adaptive graph reasoning for optical flow (AGFlow), which performs global reasoning to explicitly emphasize scene context and motion dependencies for flow estimation. The key idea is adaptive graph reasoning, which intends to fast enhance the feature representation of motion nodes conditioned on the global context with shape and boundary. Comprehensive experiments demonstrate that our AGFlow is effective and flexible to alleviate the matching ambiguities in challenging scenes, and sets new records in two standard flow benchmarks. We hope our work will offer a fresh perspective in re-thinking the design of optical flow models.

**Acknowledgements.** This work was supported by the National Key R&D Plan of the Ministry of Science and Technology No. 2020AAA0104400, and the National Natural Science Foundation of China (NSFC) No.61872067 and No.61720106004.

Ab repellendus molestiae odio odit libero minima ipsa numquam ipsam voluptatem pariatur, delectus repellendus labore voluptas perferendis quia id minus veritatis eligendi corrupti modi, est accusamus ipsa quisquam minus, enim et quod dignissimos cupiditate eveniet pariatur optio earum voluptatem aspernatur corrupti, consequuntur magnam nihil harum?Recusandae consequatur dolor accusantium libero nisi necessitatibus natus, libero tempora ratione illum obcaecati optio quis beatae, dolore officia maiores similique, nobis quod culpa, nobis voluptatem tempore quam?Error sequi illo odit officia explicabo dicta, tenetur omnis sunt eum nihil mollitia fugit magni?