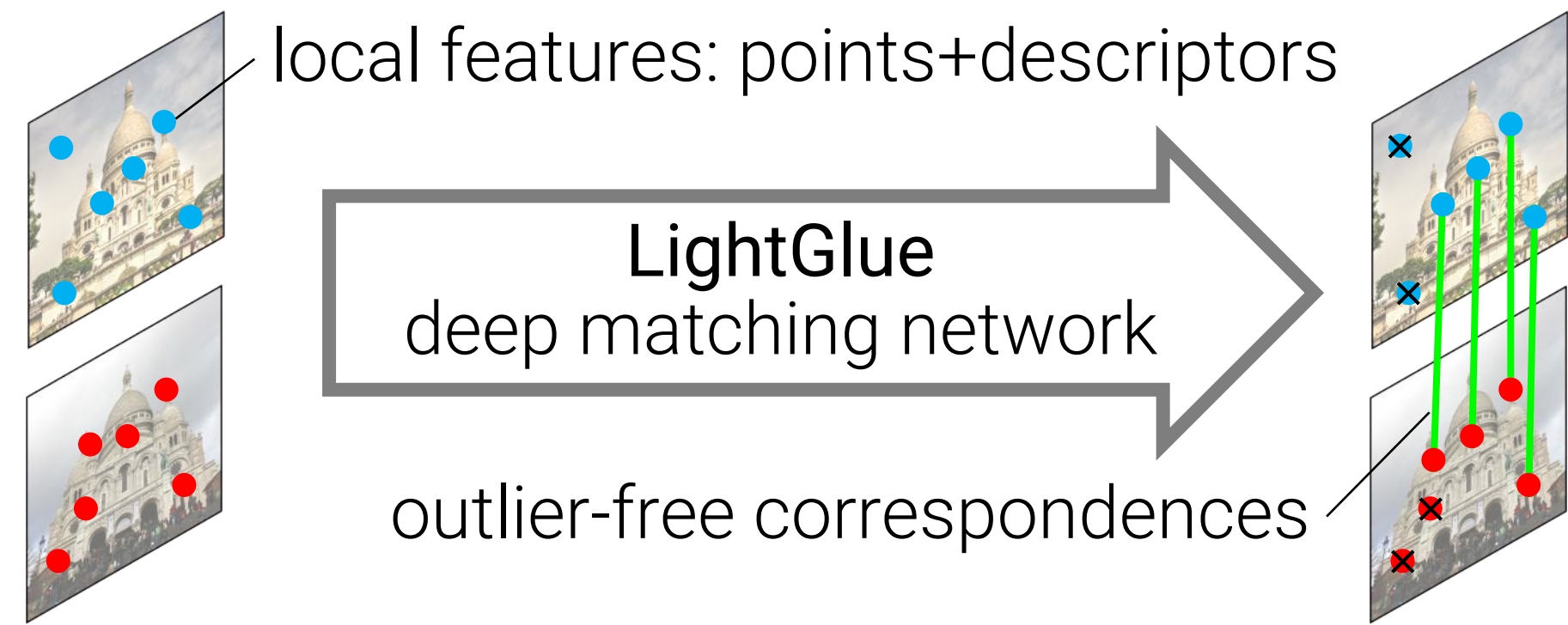
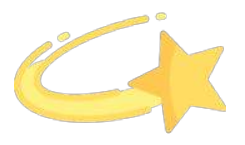
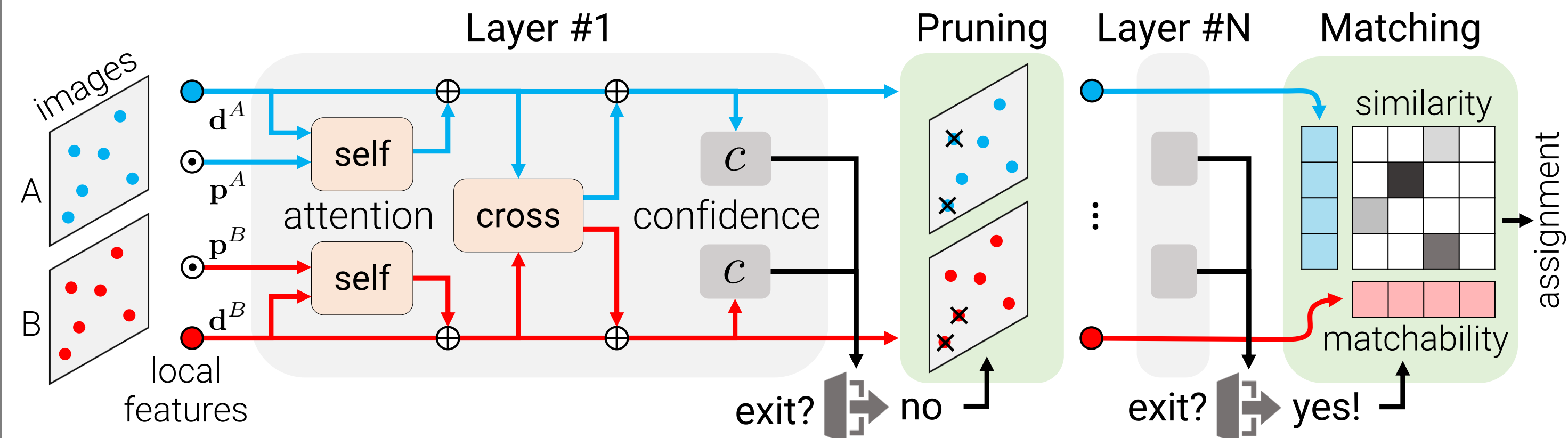


Overview



			
Accuracy	high	higher	higher
Speed	slow	fast	very slow
Training	difficult 16+ GPU-day	affordable 4-8 GPUd	difficult 40-64 GPUd

A Transformer for image matching



Key differences vs SuperGlue

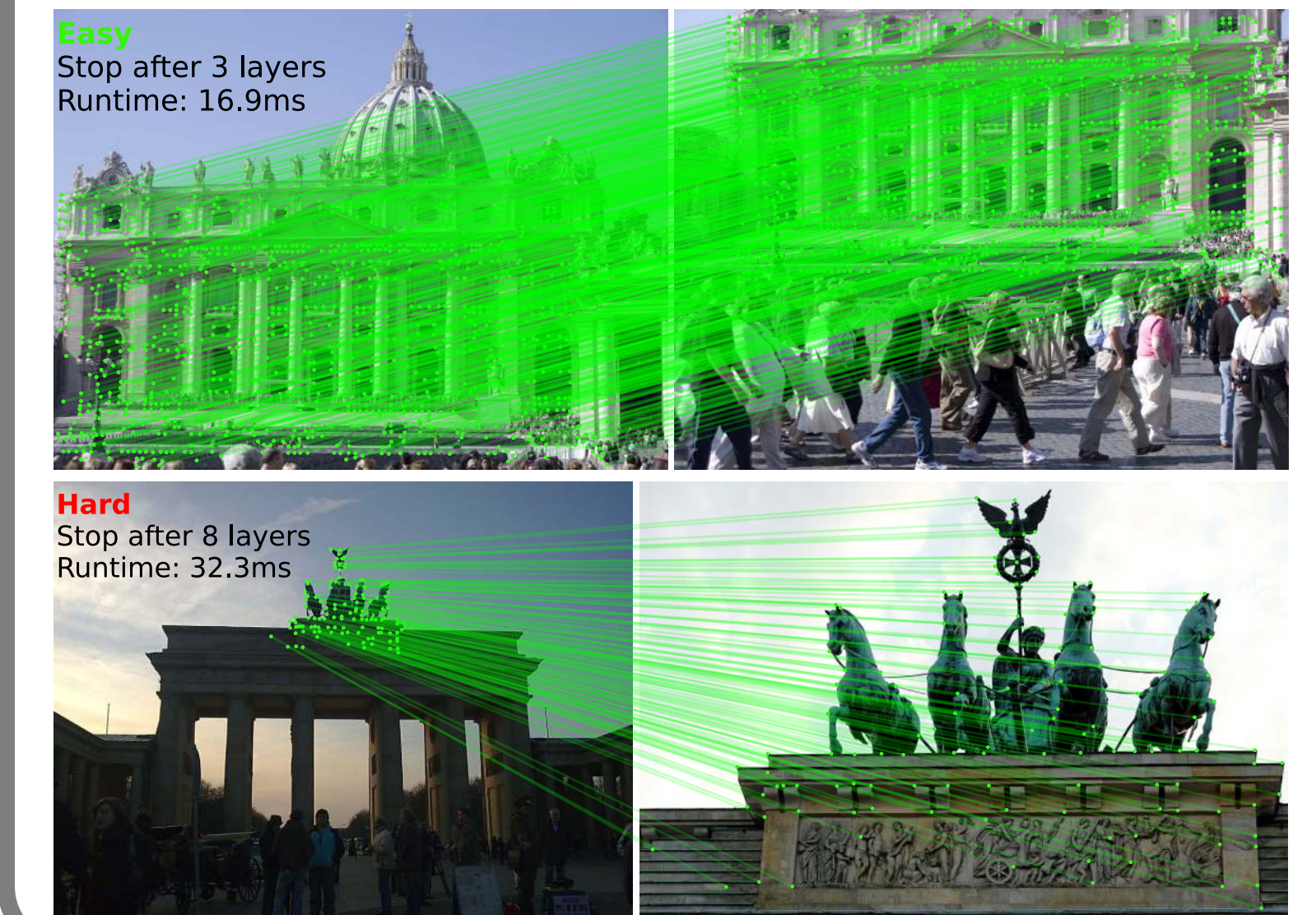
~~MLP absolute encoding~~
Relative rotary positional encoding
→ more stable training
→ higher accuracy

~~Sinkhorn assignment~~
Prediction with softmax + matchability
→ faster inference, better training

~~Fixed architecture~~
Adaptive computations in depth and width
→ faster inference

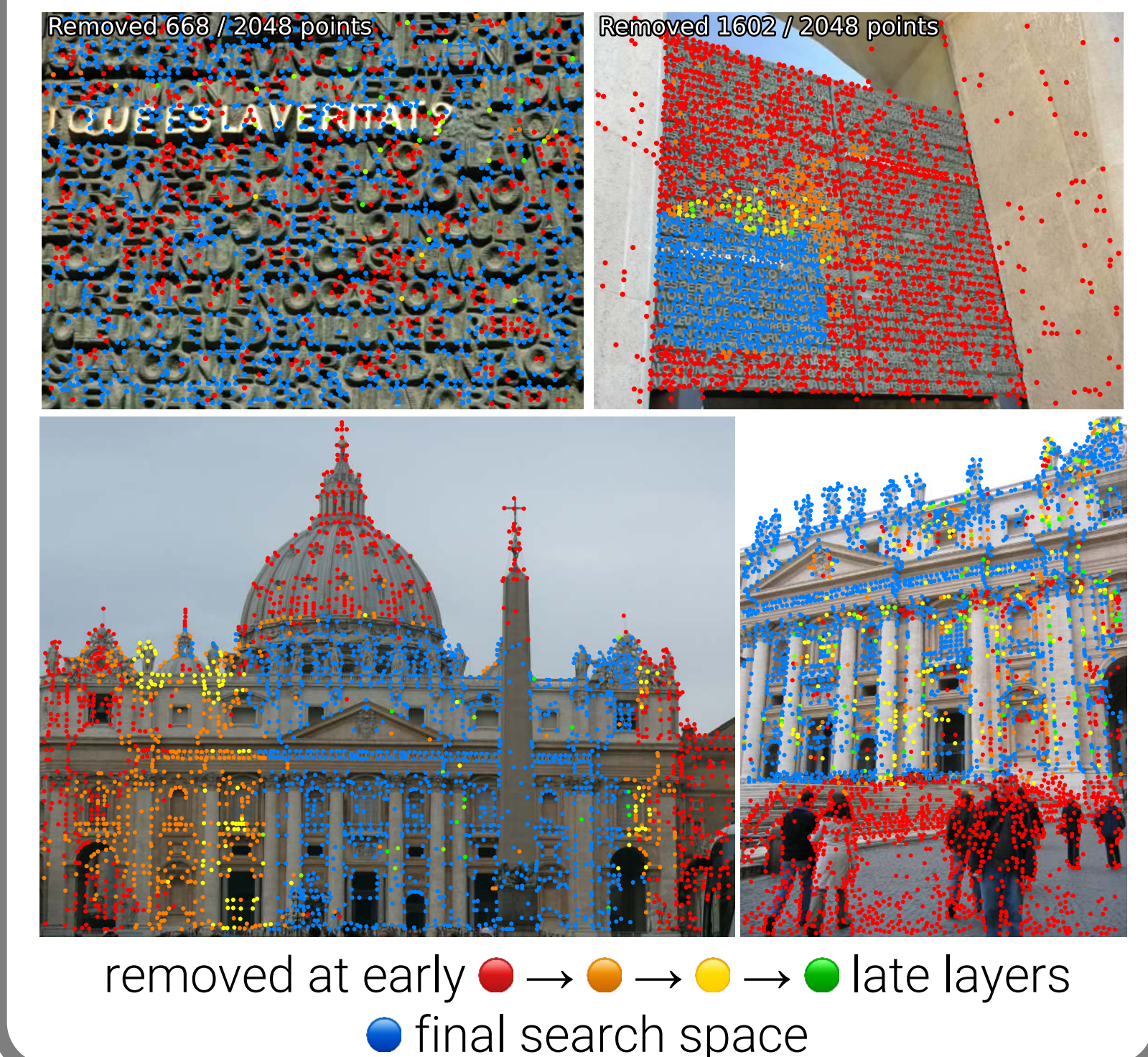
Adaptive depth

Stop at earlier layers for easier examples

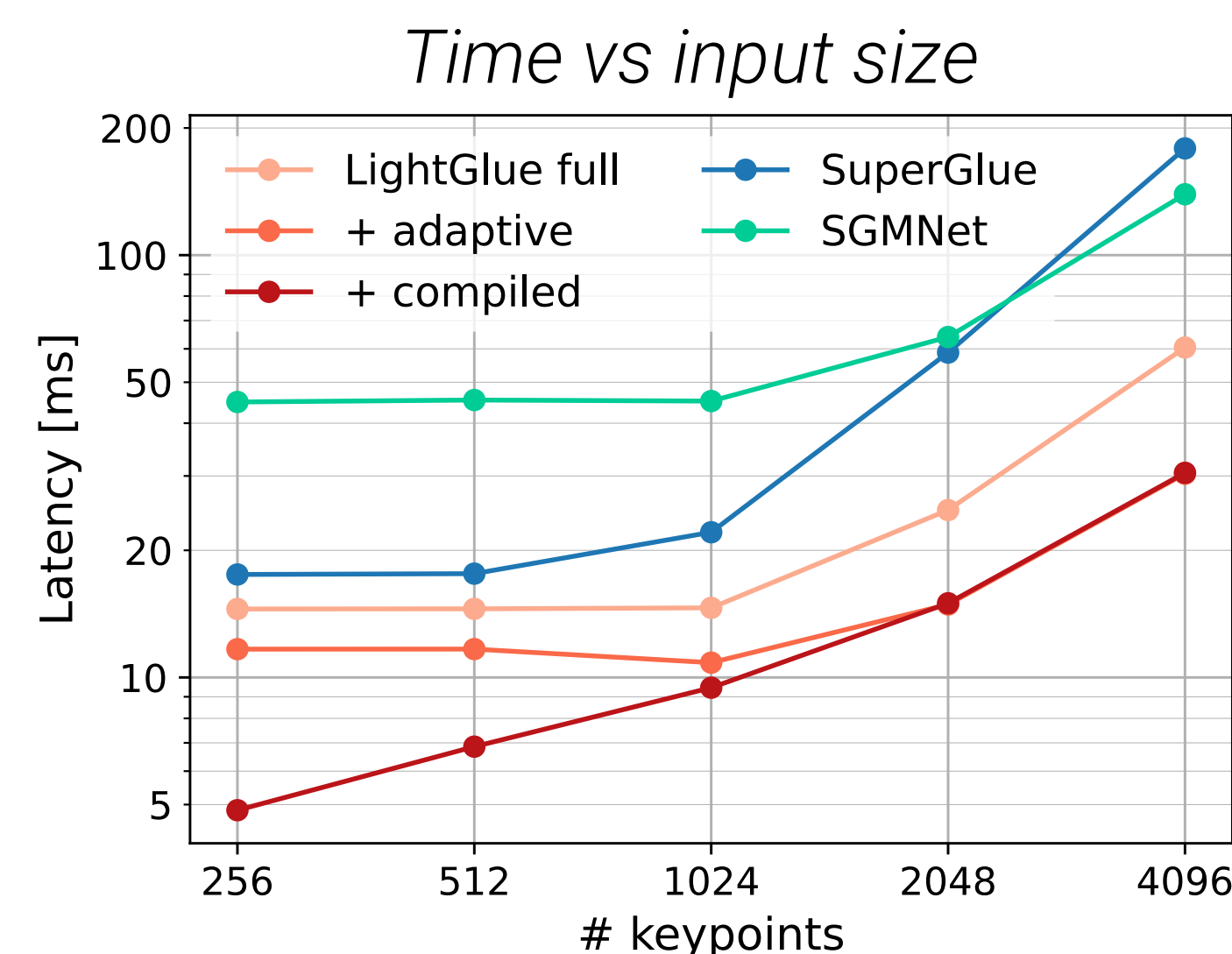
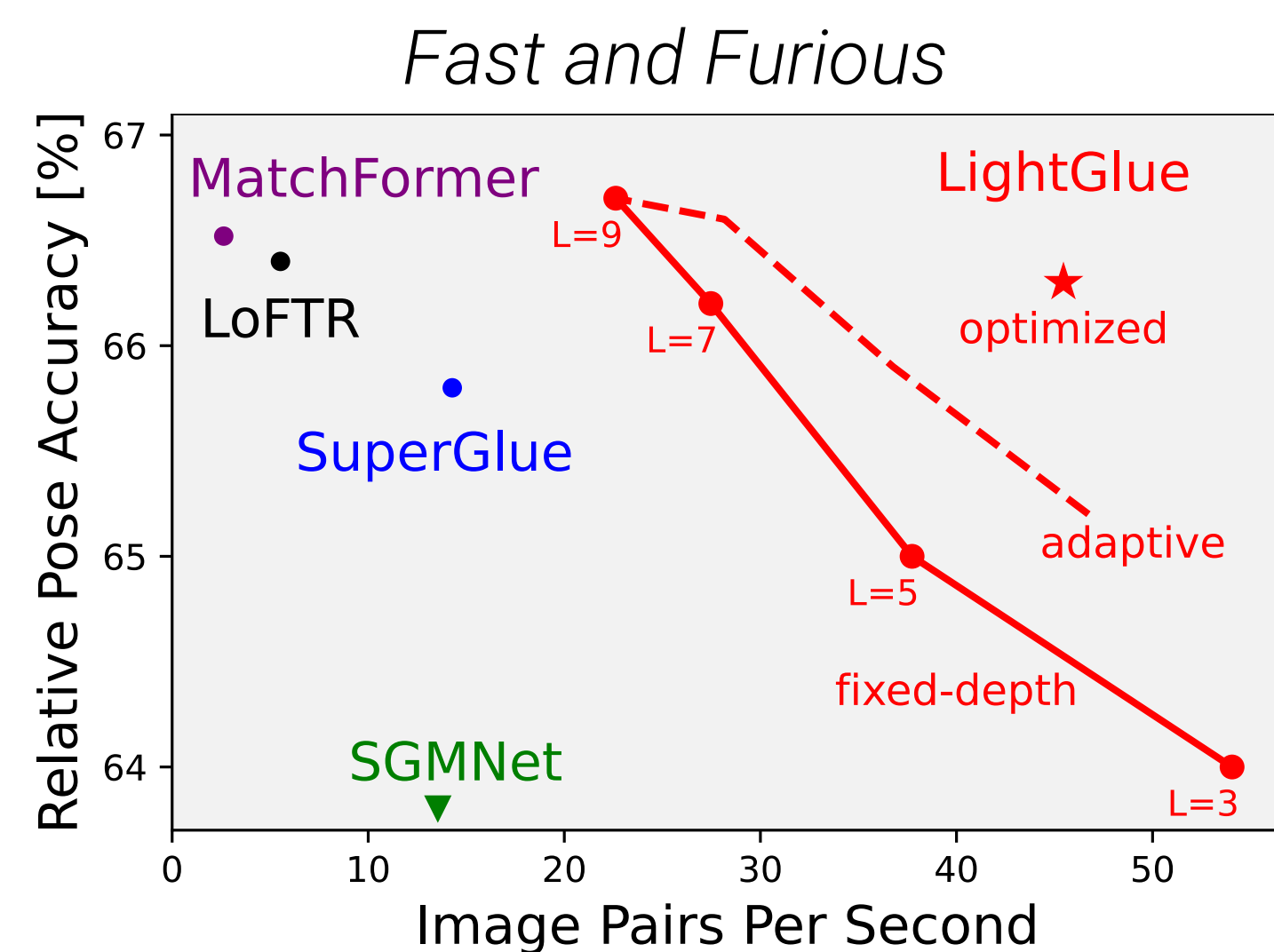


Adaptive width

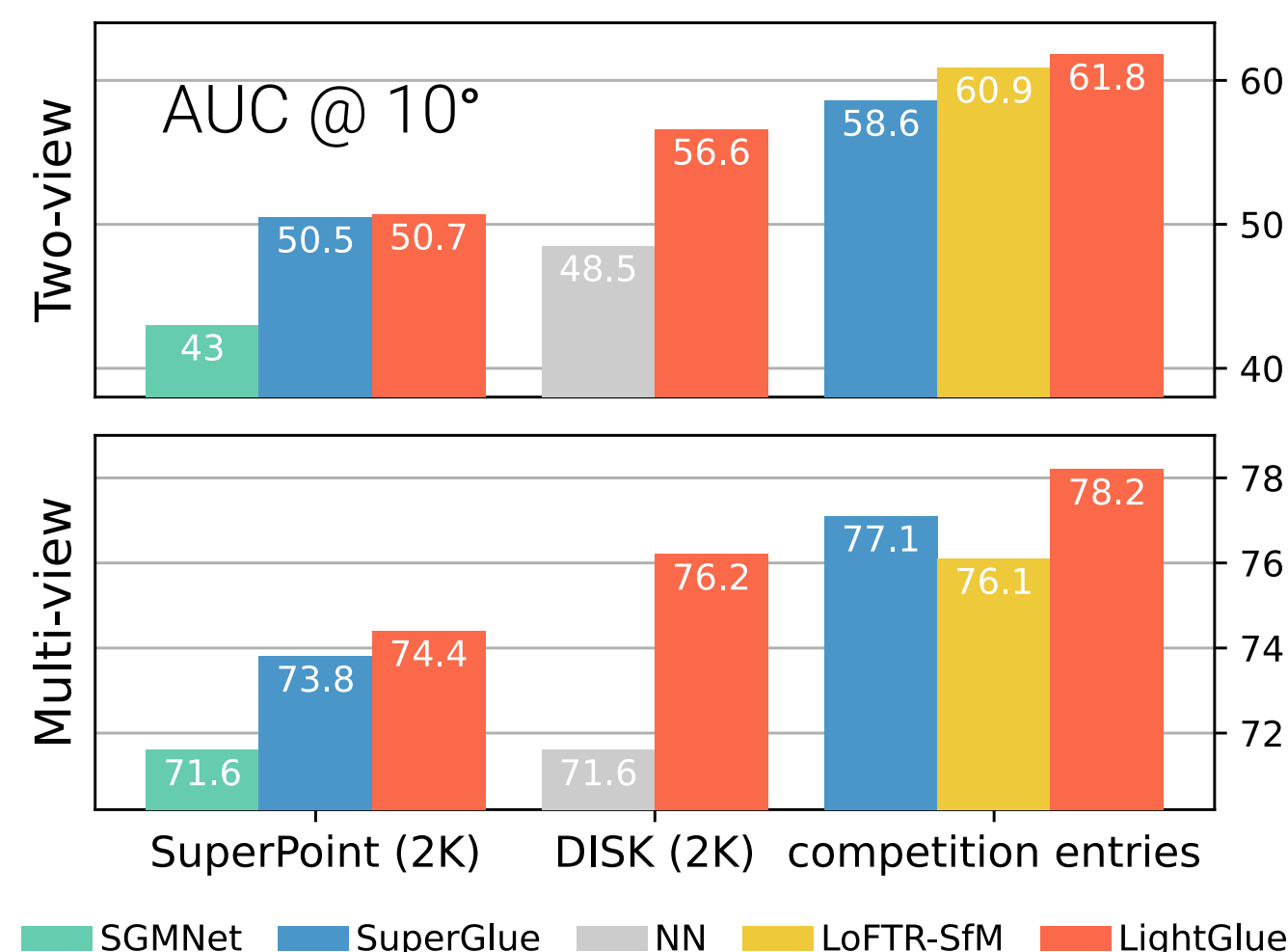
Prune points that are surely unmatched



Ready for SLAM & Structure-from-Motion



Structure-from-Motion – IMC 2021



Stopping layer

