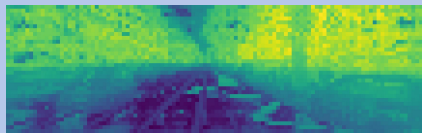


Feature Backbone for Stereo Vision

SOTA Stereo Vision Backbones

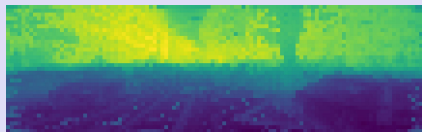


Training Without
Camera Pose

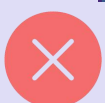


Preserve Feature

VGGT



Camera Pose
Knowledge



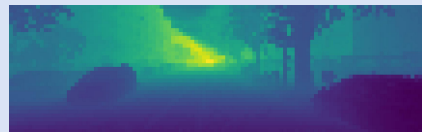
Feature-level
Structural Degradation

VS

StereoVGGT (Ours)



Camera Pose
Knowledge



Preserve
Feature

Downstream Tasks

Stereo
Conversion

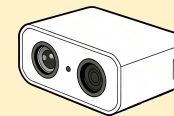


Vision Pro
et al.

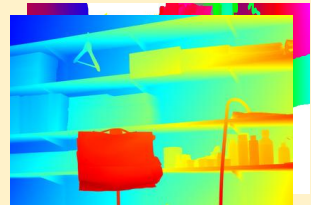


Anaglyph 3D et al.

Stereo
Matching



Binocular
Camera



Disparity Map