Temporally Consistent Online Depth Estimation in Dynamic Scenes

Zhaoshuo Li[1], Wei Ye[2], Dilin Wang[2], Francis X. Creighton[1], Russell H. Taylor[1], Ganesh Venkatesh[2], Mathias Unberath[1]

[1]Johns Hopkins University [2]Reality Labs, Meta Inc.

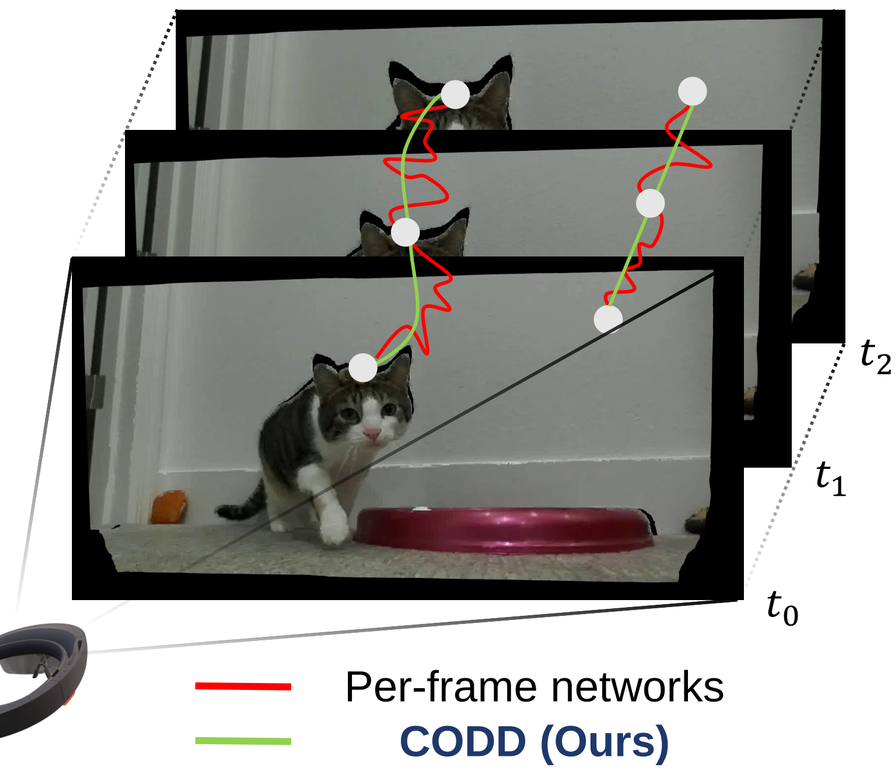
Overview

Depth estimation can be used for 3D reconstruction.

Temporally consistency has been largely overlooked.
Consistency is critical for applications such as mixed reality, as jitters in depth estimates corrupt visual quality.

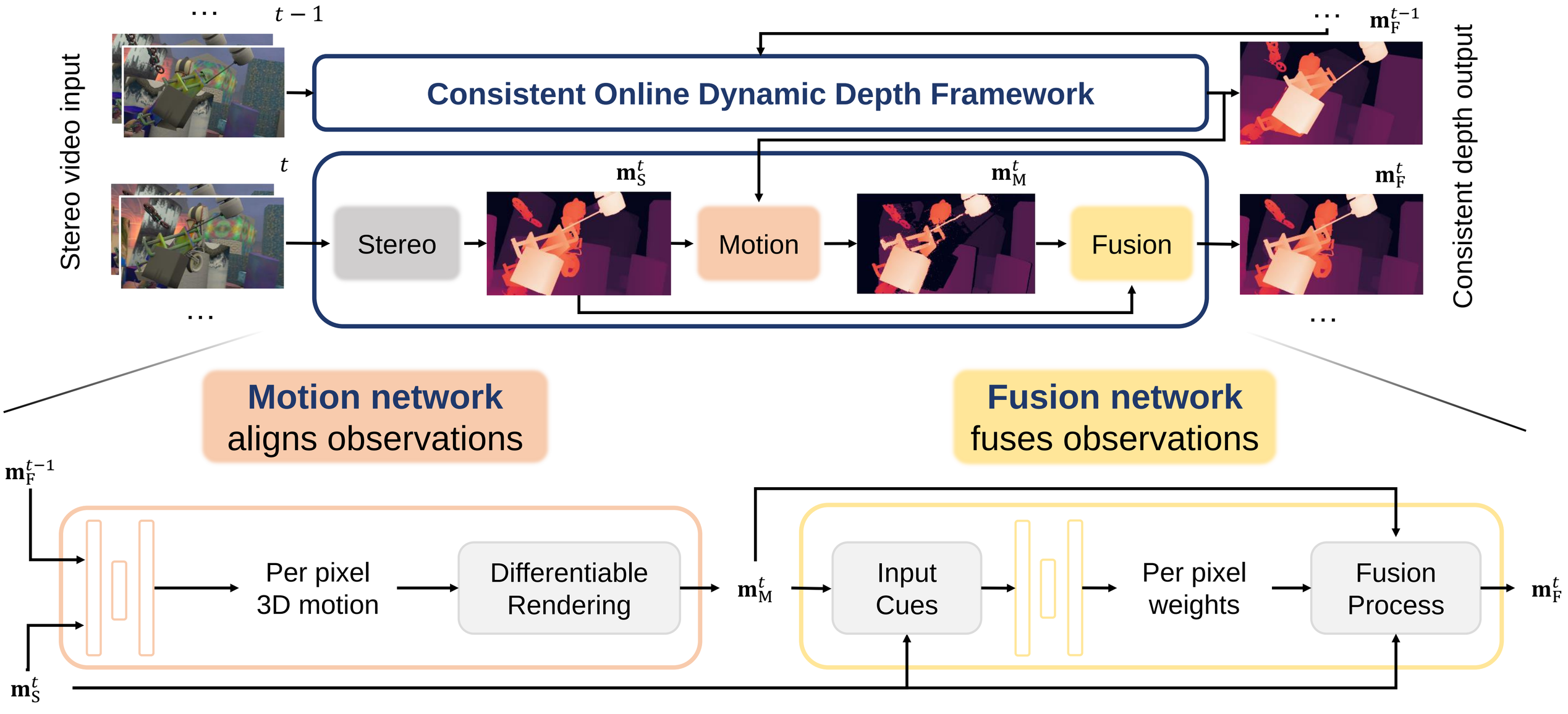
Challenges
- Online – no future information is available,
- Dynamics – camera motion, object motion and deformation.

Consistent Online Dynamic Depth (CODD) framework is developed to mitigate the above challenges.



Consistent Online Dynamic Depth Framework

CODD produces temporally **consistent** depth for **dynamic** scenes in an **online** setting.



m: memory state, per-pixel *semantic* and *disparity* estimates

Result

Metrics

EPE – per-frame depth accuracy.

$$d_{\mathrm{GT}} - d_{\mathrm{pred}}$$

TEPE – temporal depth consistency, capturing consistency of fast-moving objects.

$$\left|\left(d_{\mathrm{GT},t} - d_{\mathrm{GT},t-1}\right) - \left(d_{\mathrm{pred,t}} - d_{\mathrm{pred},t-1}\right)\right|$$

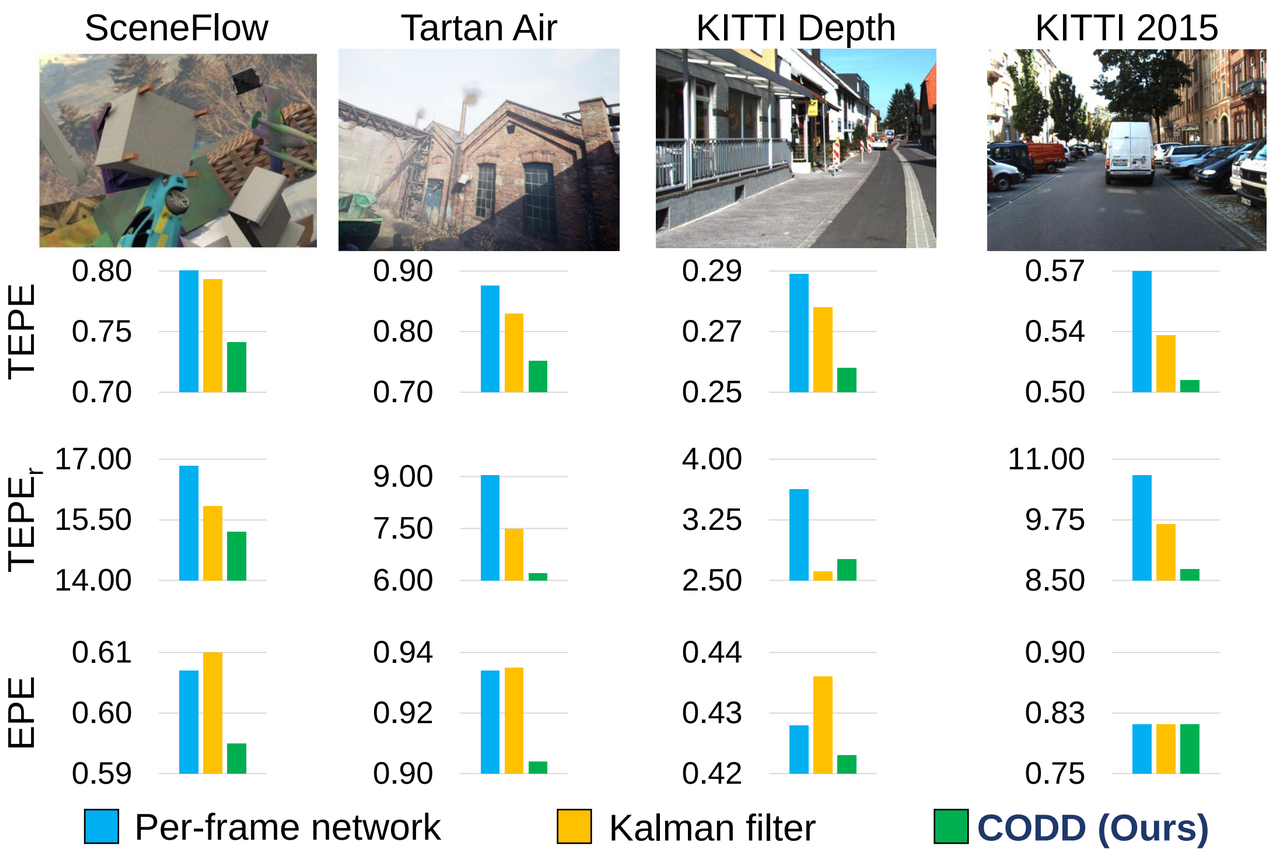
$\mathrm{TEPE}_{\mathrm{r}}$ – relative temporal depth consistency, capturing consistency of static objects.

$$\frac{\left|\left(d_{\mathrm{GT},t} - d_{\mathrm{GT},t-1}\right) - \left(d_{\mathrm{pred,t}} - d_{\mathrm{pred},t-1}\right)\right|}{\left|d_{\mathrm{GT},t} - d_{\mathrm{GT},t-1}\right|}$$

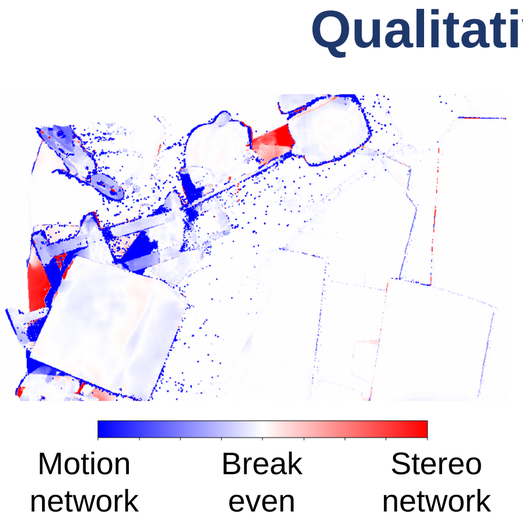
Experiments

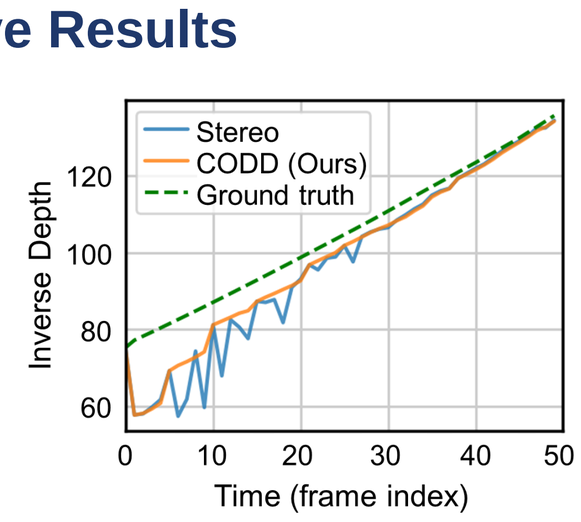
HITNet [1] is used as the per-frame stereo network. Our fusion network is compared against classical Kalman Filter algorithm [2].

CODD performs better by up to 31% for temporal depth consistency and performs on par for per-frame depth accuracy.



Qualitative Results





[1] Tankovich et al. Hitnet: Hierarchical iterative tile refinement network for real-time stereo matching. CVPR 2021.
[2] Kalman. A new approach to linear filtering and prediction problems. JFE 1960.

Checkout our project page for demos and more details!
https://mli0603.github.io/codd