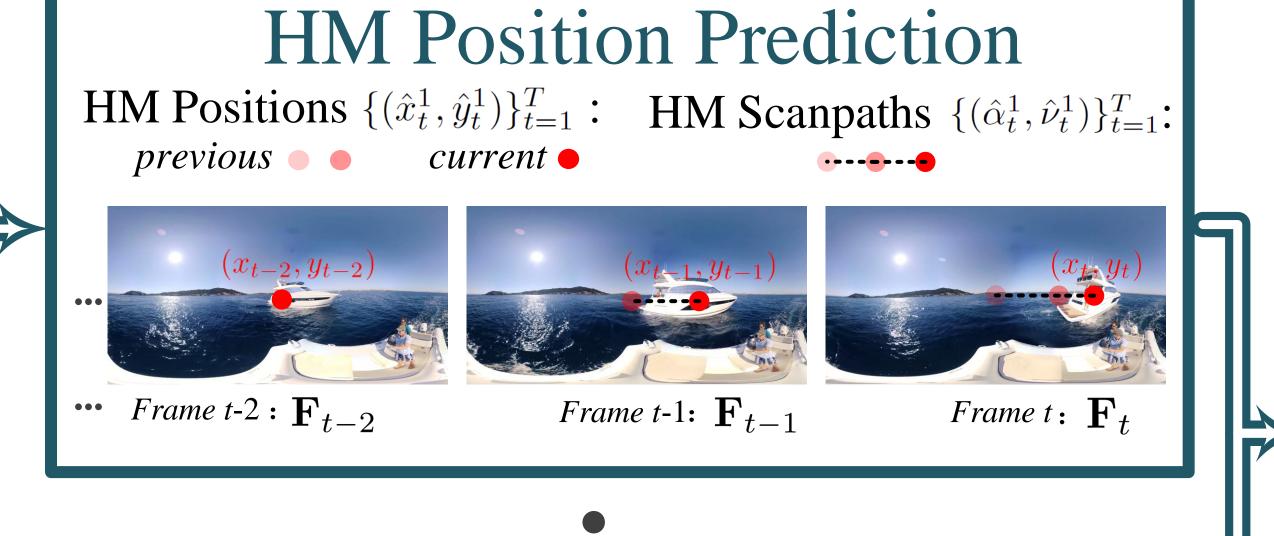
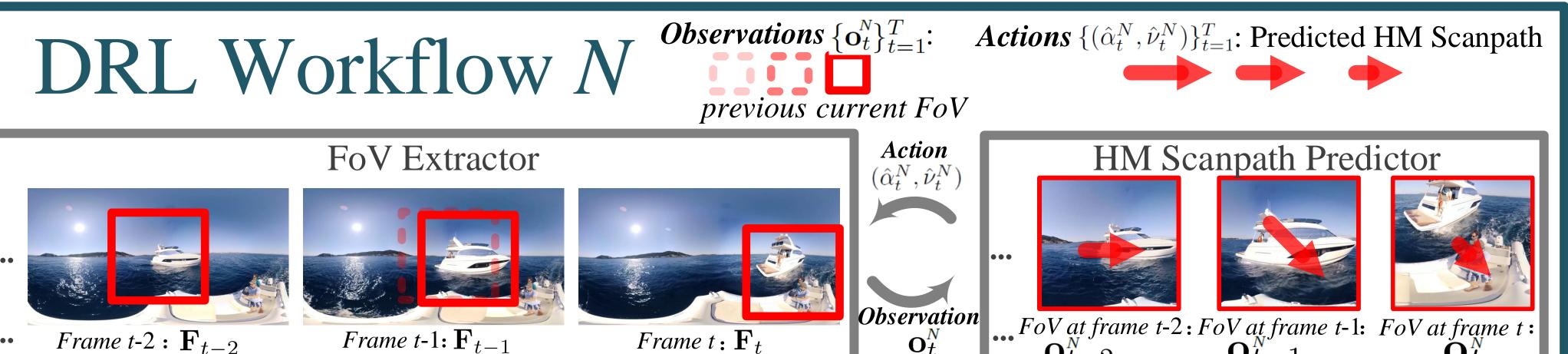
$Actions\{(\hat{\alpha}_t^1, \hat{\nu}_t^1)\}_{t=1}^T : Predicted HM Scanpath$ FoV Extractor HM Scanpath Predictor $(\hat{\alpha}_t^1, \hat{\nu}_t^1)$ **Observation** ... FoV at frame t-2: FoV at frame t-1: FoV at frame Frame t-2 : \mathbf{F}_{t-2}

Input:

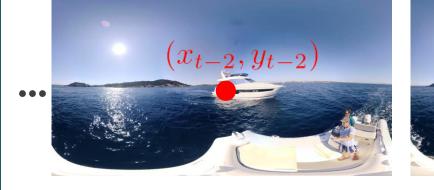
Panoramic Video

Frames $\{\mathbf{F}_t\}_{t=1}^T$













2

Frame t-2 : \mathbf{F}_{t-2}

Frame t-1: \mathbf{F}_{t-1}

Frame $t\colon \mathbf{F}_t$

Smoothing Output: \mathbf{HM} Maps \mathbf{H}_{t-2}

HM Position

Integration &