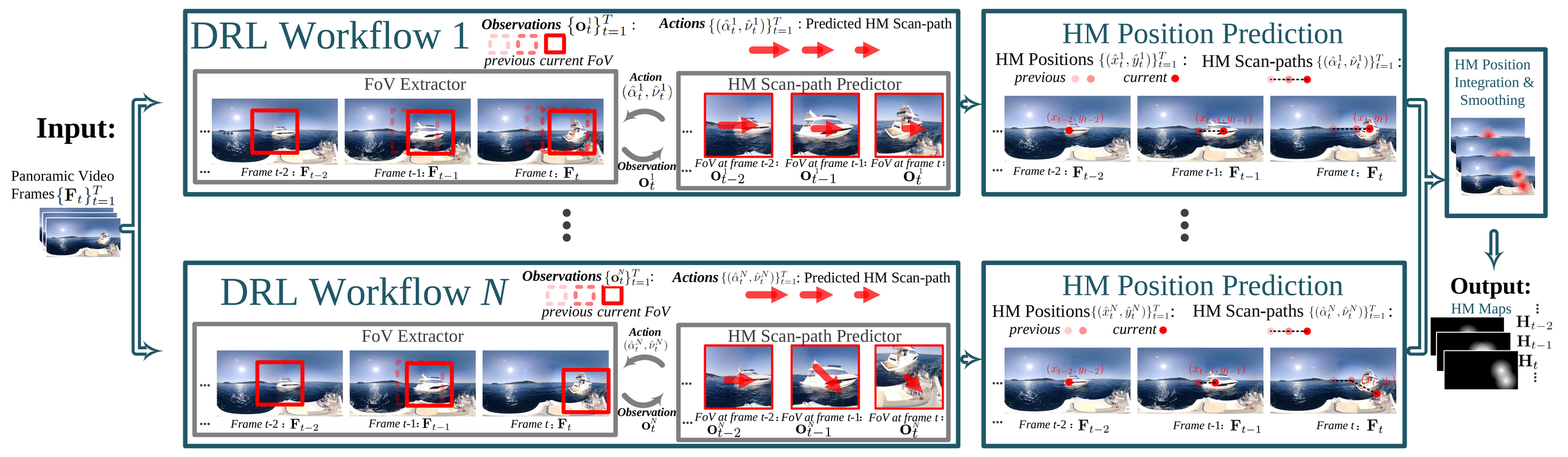
Input:
Panoramic Video Frames $\{\mathbf{F}_t\}_{t=1}^{T}$
DRL Workflow 1
Observations $\{\mathbf{o}_t^1\}_{t=1}^{T}$:
previous current FoV
Actions $\{(\hat{\alpha}_t^1, \hat{\nu}_t^1)\}_{t=1}^{T}$: Predicted HM Scan-path
FoV Extractor
Frame t-2 : $\mathbf{F}_{t-2}$
Frame t-1: $\mathbf{F}_{t-1}$
Frame t : $\mathbf{F}_t$
Action $(\hat{\alpha}_t^1, \hat{\nu}_t^1)$
Observation $\mathbf{o}_t^1$
HM Scan-path Predictor
FoV at frame t-2 : $\mathbf{o}_{t-2}^1$
FoV at frame t-1: $\mathbf{o}_{t-1}^1$
FoV at frame t : $\mathbf{o}_t^1$
HM Position Prediction
HM Positions $\{(\hat{x}_t^1, \hat{y}_t^1)\}_{t=1}^{T}$:
HM Scan-paths $\{(\hat{\alpha}_t^1, \hat{\nu}_t^1)\}_{t=1}^{T}$:
previous current
(x_{t-2}, y_{t-2})
(x_{t-1}, y_{t-1})
(x_t, y_t)
Frame t-2 : $\mathbf{F}_{t-2}$
Frame t-1: $\mathbf{F}_{t-1}$
Frame t : $\mathbf{F}_t$
DRL Workflow N
Observations $\{\mathbf{o}_t^N\}_{t=1}^{T}$:
previous current FoV
Actions $\{(\hat{\alpha}_t^N, \hat{\nu}_t^N)\}_{t=1}^{T}$: Predicted HM Scan-path
FoV Extractor
Frame t-2 : $\mathbf{F}_{t-2}$
Frame t-1: $\mathbf{F}_{t-1}$
Frame t : $\mathbf{F}_t$
Action $(\hat{\alpha}_t^N, \hat{\nu}_t^N)$
Observation $\mathbf{o}_t^N$
HM Scan-path Predictor
FoV at frame t-2 : $\mathbf{o}_{t-2}^N$
FoV at frame t-1: $\mathbf{o}_{t-1}^N$
FoV at frame t : $\mathbf{o}_t^N$
HM Position Prediction
HM Positions $\{(\hat{x}_t^N, \hat{y}_t^N)\}_{t=1}^{T}$:
HM Scan-paths $\{(\hat{\alpha}_t^N, \hat{\nu}_t^N)\}_{t=1}^{T}$:
previous current
(x_{t-2}, y_{t-2})
(x_{t-1}, y_{t-1})
(x_t, y_t)
Frame t-2 : $\mathbf{F}_{t-2}$
Frame t-1: $\mathbf{F}_{t-1}$
Frame t : $\mathbf{F}_t$
HM Position Integration & Smoothing
Output:
HM Maps
$\mathbf{H}_{t-2}$
$\mathbf{H}_{t-1}$
$\mathbf{H}_t$