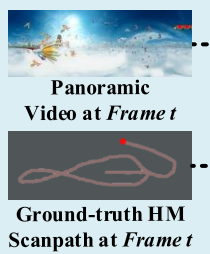
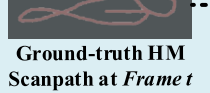


Input:

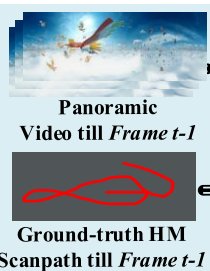
\mathbf{F}_t



$\alpha_t \nu_t$



$\{\mathbf{F}_i\}_{i=1}^{t-1}$



$\{\alpha_i\}_{i=1}^{t-1} \{\nu_i\}_{i=1}^{t-1}$

Observation

\mathbf{O}_t

Observation

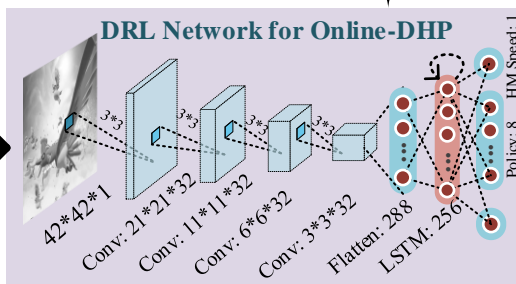
$\{\mathbf{O}_i\}_{i=1}^{t-1}$



Reward Estimator
Equation (1)(2)

Reward

$\{r_i^\alpha\}_{i=1}^{t-1} \{r_i^\nu\}_{i=1}^{t-1}$



DRL Network for Offline-DHP

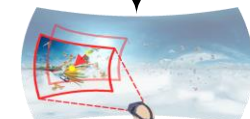
Output:

Action

$\hat{\nu}_t \hat{\alpha}_t$



+



Predicted HM Position at Frame $t+1$

$(\hat{x}_{t+1}, \hat{y}_{t+1})$

Legend



Recurrent input



Single input



Cycle operation



Single operation