Self-Guided Action Diffusion

[state encoding]

$$\mathcal{L} = \sum_{i=t+h}^{t+l} w_i \cdot \left\| \hat{a}_i - a_i^{\text{prior}} \right\|_2^2$$

d

[guided action]

x_0 → x_1 → x_2 → x_k

t_0

Prior

t_1

d_0

$\cdots$

t_1

d_k

Next Action

K steps of Iterative Denoising