

Supplementary File

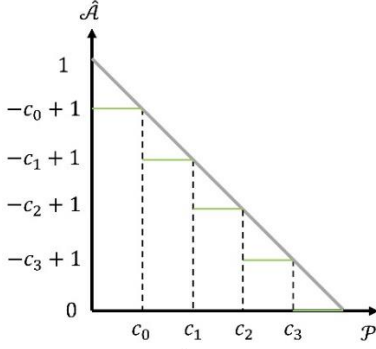
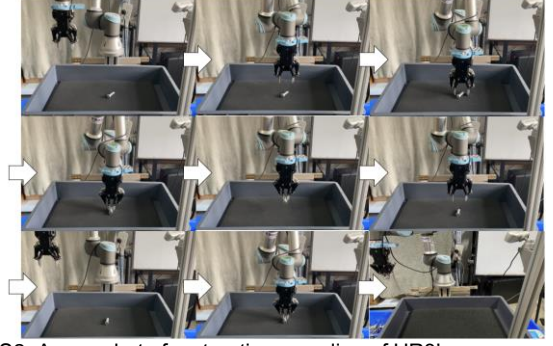
Fig. S1. Relationship between $\hat{\mathcal{A}}$ and \mathcal{P} in Eq. 8.

Fig. S2. A snapshot of restorative sampling of UR3's grasp.

Algorithm A1 Restorative Sampling

- 1: Capture I_i and obtain \mathcal{O}_i .
 - 2: Obtain \mathcal{G}_i given \mathcal{O}_i .
 - 3: // Perform restorative manipulation M_i next.
 - 4: Gripper moves first to T_i^l and rotates ϕ_i around Z-axis.
 - 5: Gripper moves to T_i along \mathcal{T}_i .
 - 6: Close Gripper.
 - 7: **if** the grasp is successful **then**
 - 8: Gripper moves backward to T_i^l along \mathcal{T}_i^- .
 - 9: **if** the object is held during the return **then**
 - 10: Places the object back to T_i along \mathcal{T}_i .
 - 11: Set $g_i = 1$, capture \mathcal{O}_i^+ , and calculate \mathcal{S}_i .
 - 12: Perform \mathcal{G}_i again to take the object to its goal place.
 - 13: **else**
 - 14: Set $g_i = 0$ and $\mathcal{S}_i = 0$.
 - 15: **end if**
 - 16: **else**
 - 17: Set $g_i = 0$ and $\mathcal{S}_i = 0$.
 - 18: **end if**
 - 19: Robot goes backs to its home.
 - 20: Store $(\mathcal{O}_i, \mathcal{G}_i, \mathcal{S}_i)$ into D .
-

Algorithm A2 Training of FAGL

- 1: Initialize RA-Net Q_θ and Target RA-Net Q_{θ^-} .
 - 2: Set hyperparameters $\alpha = 10^{-3}$, $B=16$, $step_{max} = 2500$, $\epsilon = 0.5$, $\tau = 3$, $step = 0$, and $D = \emptyset$.
 - 3: **while** $step < step_{max}$ **do**
 - 4: Obtain \mathcal{O}_i .
 - 5: $\epsilon = \text{explore_schedule}()$, $p = \text{rand}()$.
 - 6: $\mathcal{G}_i = \begin{cases} \text{argmax}_{\mathcal{G}} Q_\theta(\mathcal{O}, \mathcal{G}) & \text{if } p \leq 1 - \epsilon \\ \text{random policy} & \text{if } p > \epsilon \end{cases}$
 - 7: Obtain \mathcal{A}_i , g_i , and $\mathcal{S}_i(\hat{\mathcal{A}}_i, g_i)$.
 - 8: $D = D \cup \{(\mathcal{O}_i, \mathcal{G}_i, \mathcal{S}_i)\}$.
 - 9: $step = step + 1$.
 - 10: **if** $|D| > B$ **then**
 - 11: Random sample $\{(\mathcal{O}_i, \mathcal{G}_i, \mathcal{S}_i)\}_{i \in [0, B]}$ in D .
 - 12: Update $Q_\theta(\mathcal{O}, \mathcal{G})$ on $\{(\mathcal{O}_i, \mathcal{G}_i, \mathcal{S}_i)\}_{i \in [0, B]}$.
 - 13: **end if**
 - 14: **if** $step \% \tau == 0$ **then**
 - 15: $\theta^- = \theta$.
 - 16: **end if**
 - 17: **end while**
 - 18: $Q_\theta^*(\mathcal{O}, \mathcal{G}) = Q_\theta(\mathcal{O}, \mathcal{G})$.
 - 19: Output: optimal action-value function $Q_\theta^*(\mathcal{O}, \mathcal{G})$.
-

TABLE S-I
ABBREVIATIONS

\mathcal{P}	Destructive effect of a grasp (DEG)
\mathcal{G}	A grasp
$T = (x, y, z)$	The position of a grasp
ϕ	Rotation around Z-axis
ω	Initial distance between the two fingers of the gripper
I_{hc}^-	Color heightmap of the environment before restorative manipulation
I_{hc}^+	Color heightmap of the environment after restorative manipulation
\mathcal{M}	A restorative manipulation
\mathcal{B}	OTSU operation
\mathcal{A}	Antipodal degree of a grasp (ADG)
\mathcal{O}	Observation space
\mathcal{G}	Robotic grasp space
\mathcal{R}	Feedback (rewards)
ρ	Transition probability
\mathcal{H}	tMaximum step
$\hat{\mathcal{A}}$	Discrete ADG
\mathcal{T}	Trajectory
\mathcal{T}^-	Inverse trajectory of \mathcal{T}
H	Weight of I_{hc}^- and I_{hc}^+
W	Height of I_{hc}^- and I_{hc}^+
I	RGB-D image
I_c	RGB image
I_d	Depth map
g	Grasp flag
\mathcal{C}	Destruction tolerance
r	:Primary basic reward
e	Secondary basic reward
\mathcal{Q}	Multi-channel grasp affordance
q	Affordance planes
I_{hci}	Color heightmap images
I_{hdi}	Depth heightmap images
\mathcal{V}_{sd}	Shallow depth feature
\mathcal{V}_{sc}	Shallow color feature
\mathcal{V}_m	Latent feature