Supplementary File

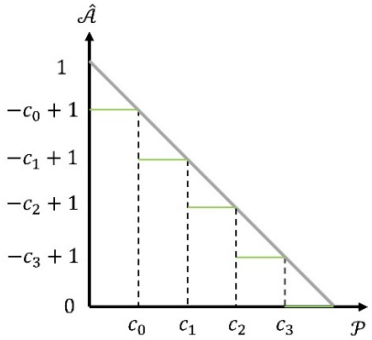

Fig. S1. Relationship between $\tilde{\mathcal{A}}$ and $\mathcal{P}$ in Eq. 8.

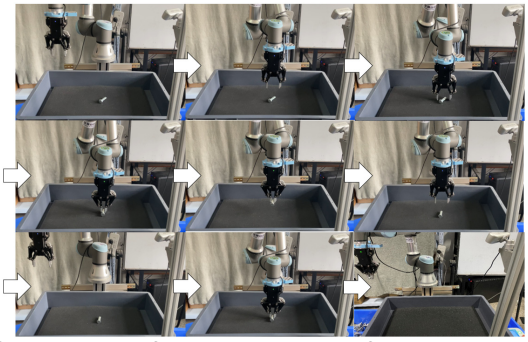

Fig. S2. A snapshot of restorative sampling of UR3's grasp.

Algorithm A1 Restorative Sampling

Capture I_i and obtain $\mathcal{O}_i$.
Obtain $\mathcal{G}_i$ given $\mathcal{O}_i$.
// Perform restorative manipulation M_i next.
Gripper moves first to T_i^l and rotates ϕ_i around Z-axis.
Gripper moves to T_i along $\mathcal{T}_i$.
Close Gripper.
if the grasp is successful **then**
  Gripper moves backward to T_i^l along $\mathcal{T}_i^-$.
  if the object is held during the return **then**
    Places the object back to T_i along $\mathcal{T}_i$.
    Set $g_i = 1$, capture $\mathcal{O}_i^+$, and calculate $\mathcal{S}_i$.
    Perform $\mathcal{G}_i$ again to take the object to its goal place.
  else
    Set $g_i = 0$ and $\mathcal{S}_i = 0$.
  end if
else
  Set $g_i = 0$ and $\mathcal{S}_i = 0$.
end if
Robot goes backs to its home.
Store $(\mathcal{O}_i, \mathcal{G}_i, \mathcal{S}_i)$ into D.

Algorithm A2 Training of FAGL

Initialize RA-Net Q_θ and Target RA-Net Q_{θ^-} .
Set hyperparameters $\alpha = 10^{-3}$, B=16, $step_{max} =$ 2500, ϵ = 0.5, $\tau = 3$, $step = 0$, and $D = \emptyset$.
while $step < step_{max}$ **do**
  Obtain $\mathcal{O}_i$.
  $\epsilon = explore\_schedule()$, $p = rand()$.
  $\mathcal{G}_i = \begin{cases} \underset{\mathcal{G}}{argmax} Q_\theta(\mathcal{O}, \mathcal{G}) & if\ p \le 1-\epsilon \\ random\ policy & if\ p > \epsilon \end{cases}$
  Obtain $\mathcal{A}_i$, g_i, and $\mathcal{S}_i(\hat{\mathcal{A}}_i, g_i)$.
  $D = D \cup \{(\mathcal{O}_i, \mathcal{G}_i, \mathcal{S}_i)\}$.
  $step = step + 1$.
  if $|D| > B$ **then**
    Random sample $\{(\mathcal{O}_i, \mathcal{G}_i, \mathcal{S}_i)\}_{i\in[0,B]}$ in D.
    Update $Q_\theta(\mathcal{O}, \mathcal{G})$ on $\{(\mathcal{O}_i, \mathcal{G}_i, \mathcal{S}_i)\}_{i\in[0,B]}$.
  end if
  if $step$ % τ == 0 **then**
    $\theta^- = \theta$.
  end if
  end while
$Q_\theta^*(\mathcal{O}, \mathcal{G}) = Q_\theta(\mathcal{O}, \mathcal{G})$.
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_{hc_i}	Color heightmap images
I_{hd_i}	Depth heightmap images
$\mathcal{V}_{sd}$	Shallow depth feature
$\mathcal{V}_{sc}$	Shallow color feature
$\mathcal{V}_m$	Latent feature