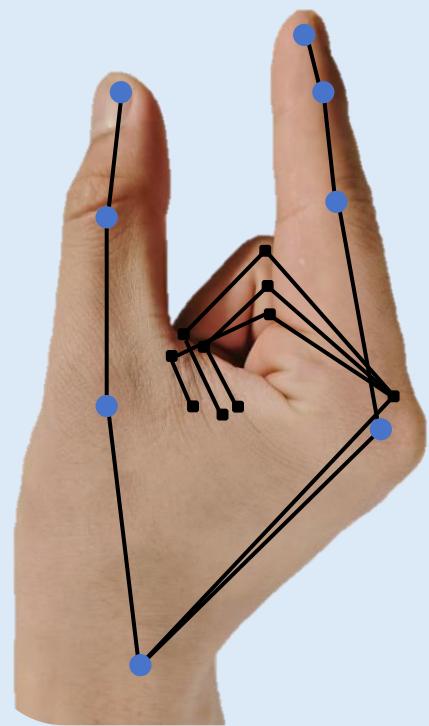


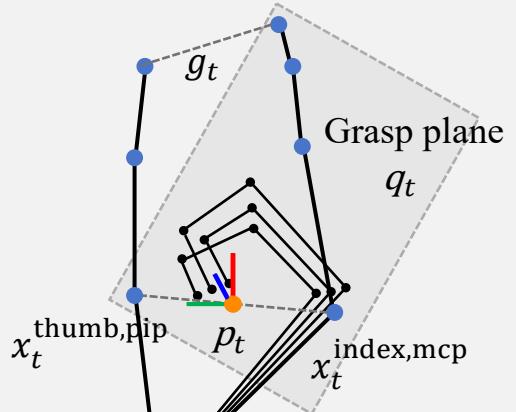
Physical Alignment



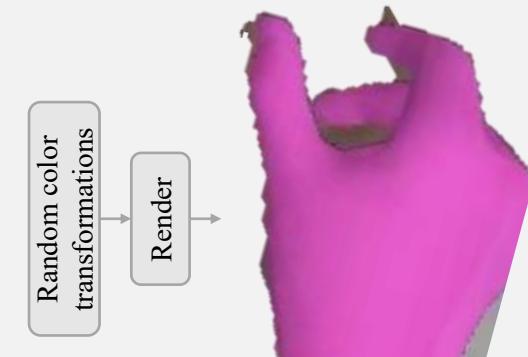
Human hand image

X_t

M_t

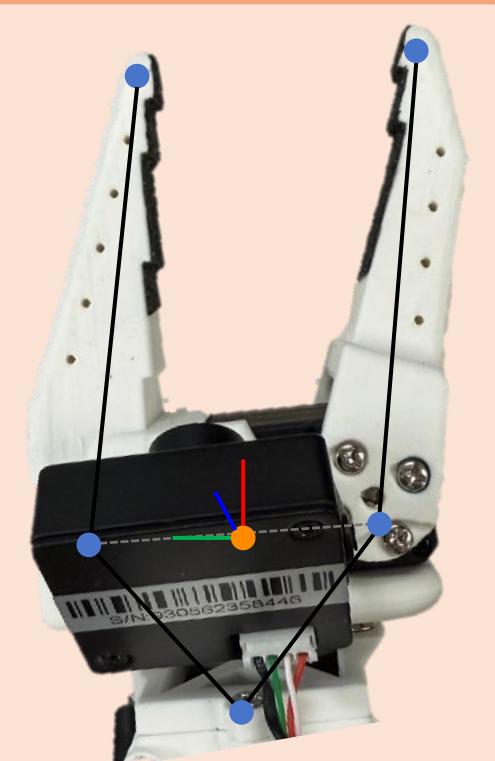


Action-space Alignment (AA)



Visual-space Alignment (VA)

P_t, I_t^{aug}



Real robot gripper