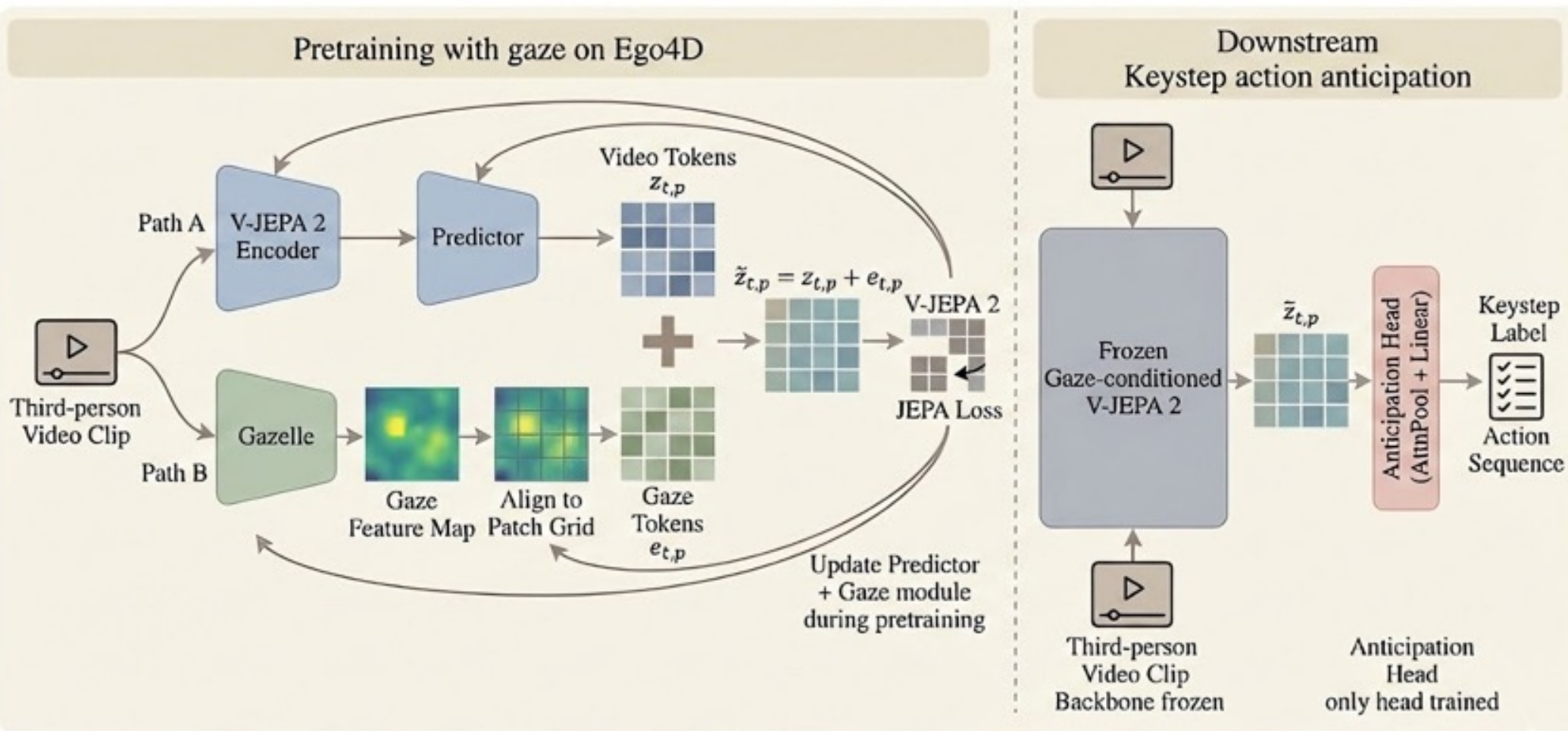


Figure 1: **Overall Framework / Two-stage World Model** → Human World Model



Overview of our gaze-augmented human world model built on V-JEPA~2.