Upsample & Aggregate   **Refine**

Feature Map

Token Map

Multi-Scale
Encode

Scale 1 l l m m K

Next-Scale Pred.
(Scale: $1 \rightarrow l$)

Scale-Travel
(Scale: $l \rightarrow m$)

Next-Scale Pred.
(Scale: $m \rightarrow K$)