

## IMMEDIATE

1



## Dynamic Visual Space (4D)

DA3's ray formulation **naturally extends to MOTION RAYS**:  $r(t) = (t(t), d(t))$ . Enables per-pixel trajectory encoding for scene flow.

Scene Flow

Video

Dynamic Nerf

## NEAR TERM

2



## Uncertainty &amp; Calibration

DA3 already predicts **depth confidence**  $D_c$ . Extend to **RAY CONFIDENCE** for robust pose alignment under occlusion.

Probabilistic

Active Vision

Safety

## MID TERM

3



## Efficiency &amp; Real-Time

Address DA3's  $O(N_v \cdot H \cdot W)^2$  **cross-view cost** with token pruning and sparse attention patterns for  $L_g$  layers.

Edge AI

Sparsity

Latency

## LONG TERM

4



## Semantic &amp; Task Coupling

Integrate language priors into **ray prediction**. Add **differentiable Bundle Adjustment** for end-to-end refinement of DA3 outputs.

Semantics

Diff. BA

Self-Supervised



**Research Focus:** Combine **Self-Supervised Cycle Consistency** with **Uncertainty Estimation** using DA3 backbone for robust, label-free learning.