EAGLE : <u>E</u>fficient <u>A</u>daptive <u>G</u>eometry-based <u>Le</u>arning in Cross-view Understanding

Thanh-Dat Truong, Utsav Prabhu, Dongyi Wang

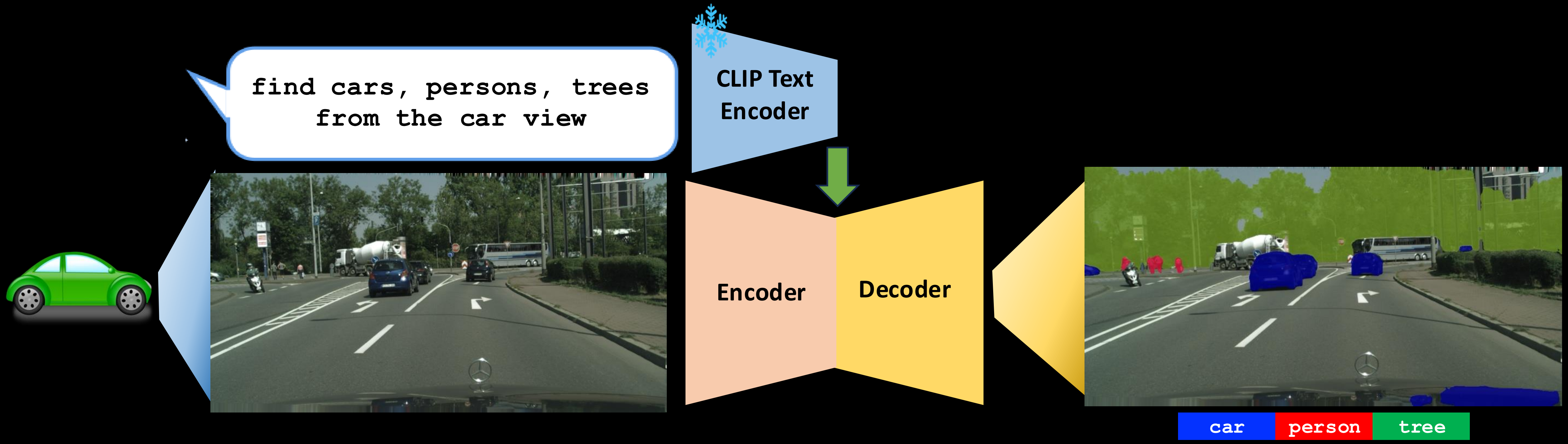
Bhiksha Raj, Susan Gauch, Jeyamkondan Subbiah, Khoa Luu

CVIU Lab, University of Arkansas Google DeepMind

University of Arkansas Carnegie Mellon University MBZUAI

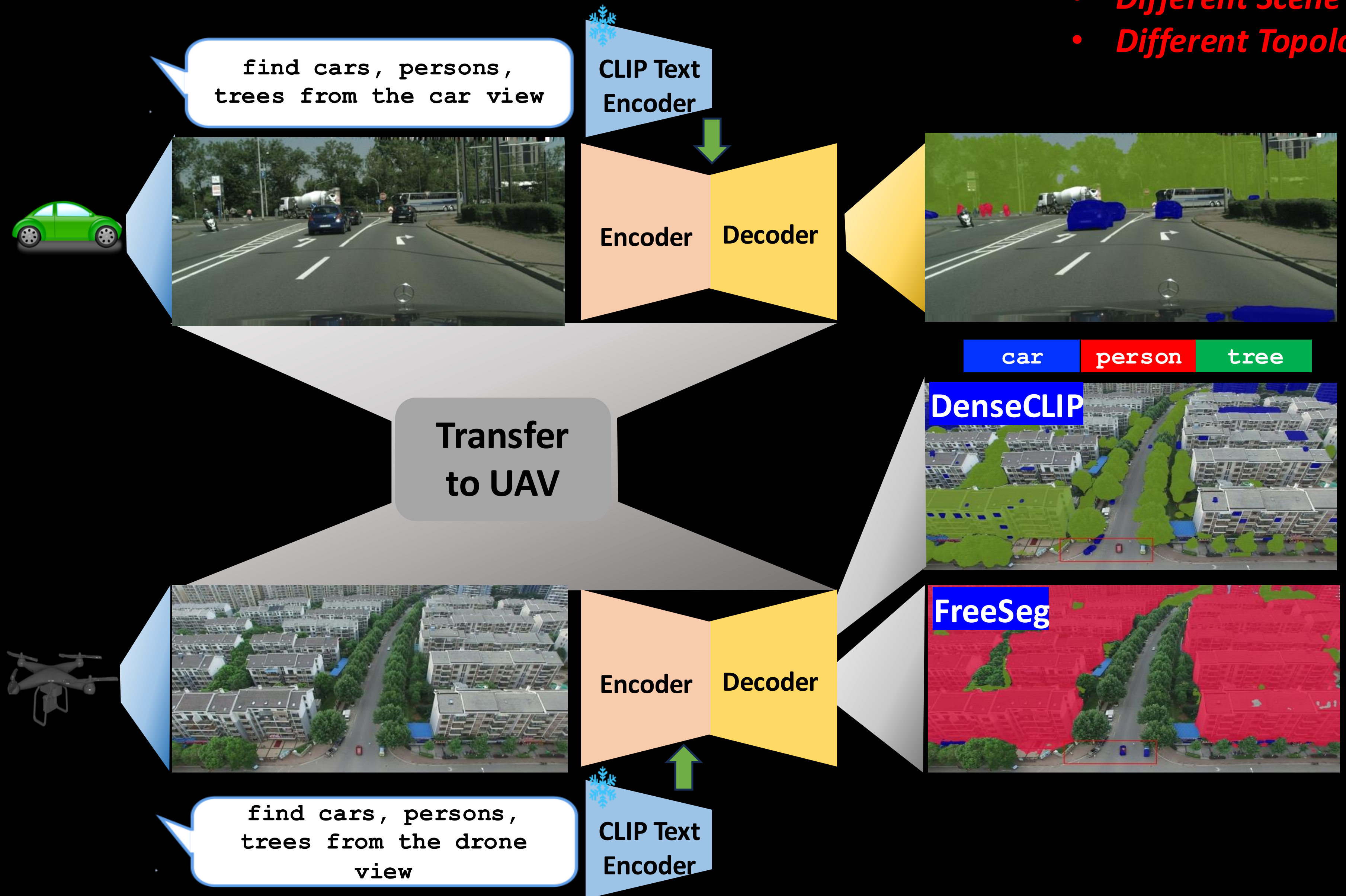
https://uark-cviu.github.io/projects/EAGLE/

Cross-view Adaptation



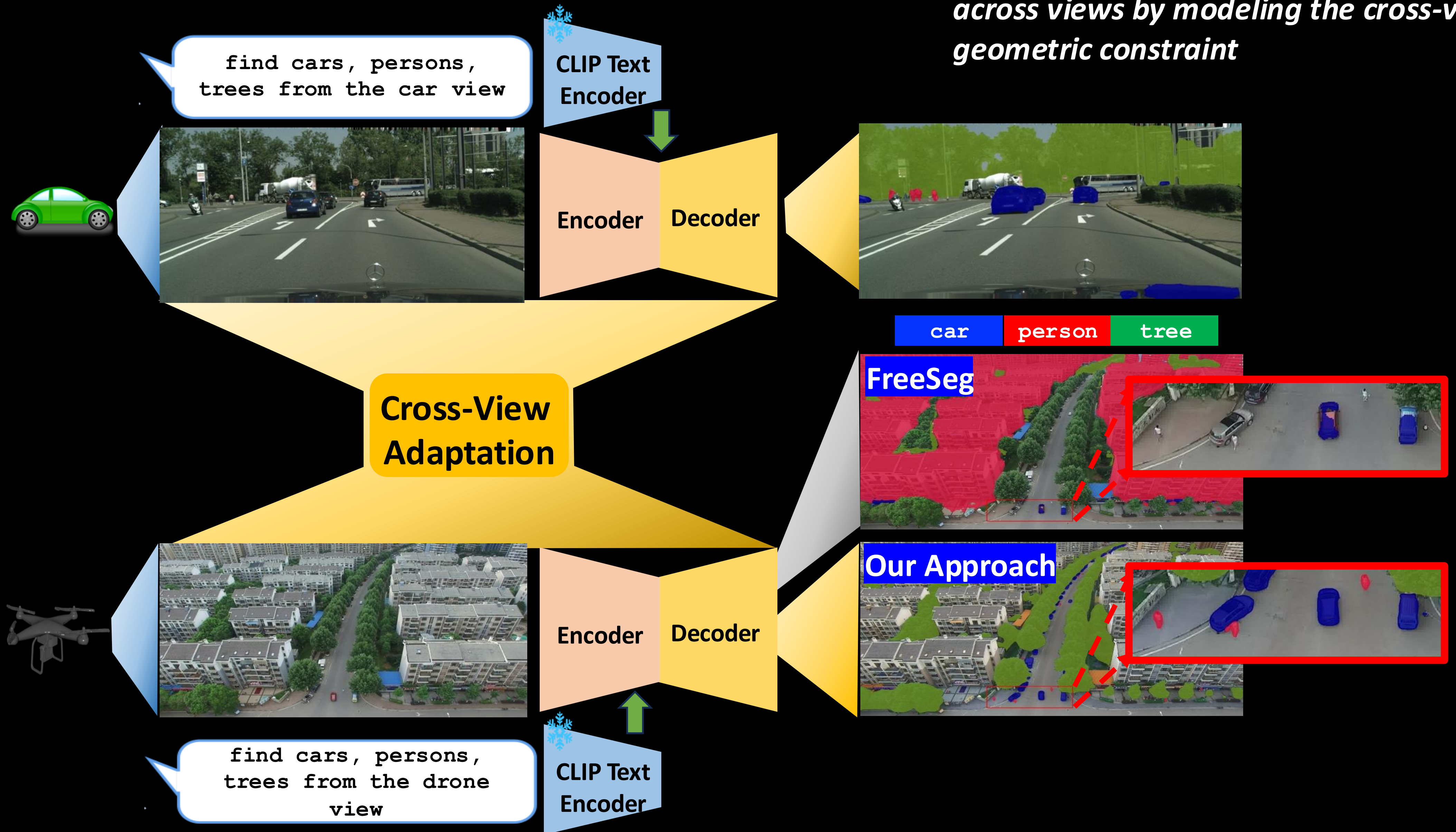
Cross-view Adaptation

Worse Performance due to
- *Different Camera Position*
- *Different Scene Structures*
- *Different Topological Layout*



Cross-view Adaptation

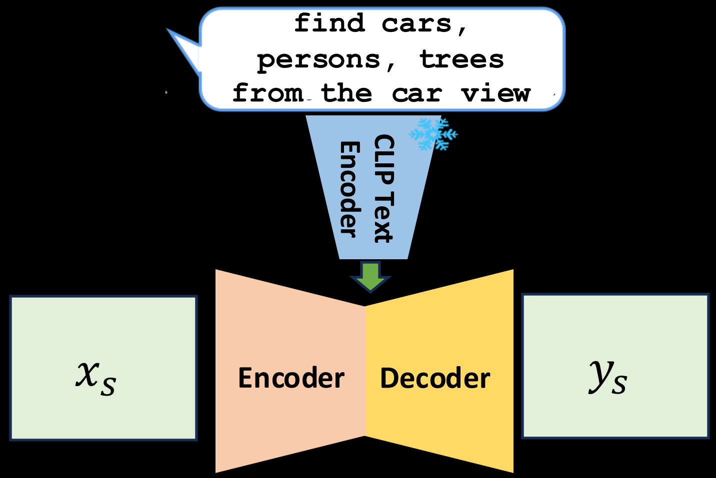
✅ *Our Cross-view Adaptation Approach improves the performance of segmentation across views by modeling the cross-view geometric constraint*



Contributions

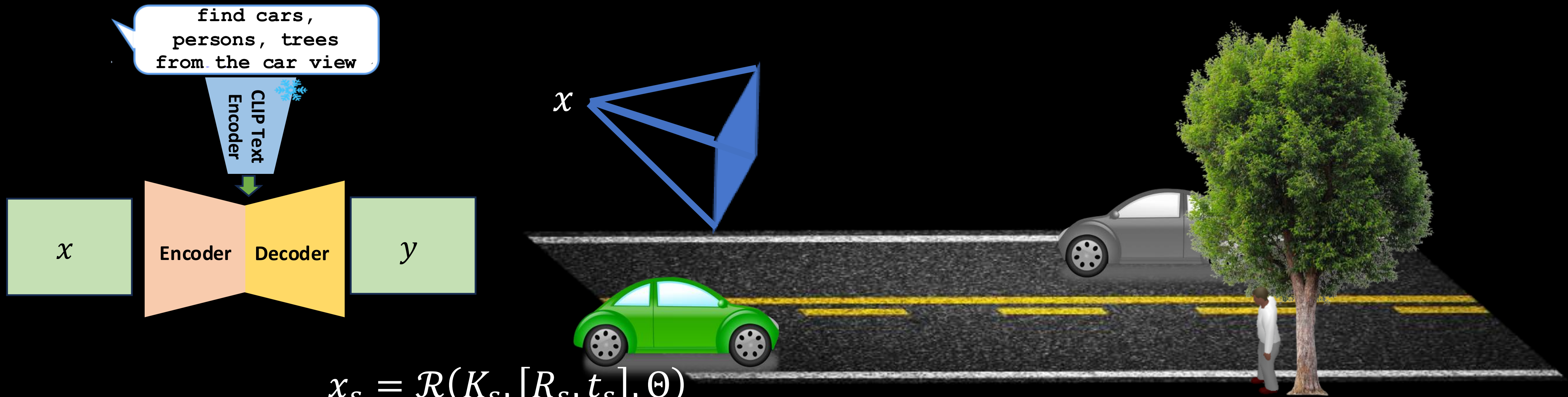
- ✓ Introduce a novel ***Cross-view Adaptation Learning Approach*** to ***Semantic Segmentation***
  - ✓ Present a ***new cross-view geometric constraints*** between image and segmentation spaces
  - ✓ Propose a new ***Geodesic Flow-based metric*** to measure the ***cross-view structural changes***
  - ✓ Introduce ***a novel view-condition prompting mechanism*** to cross-view adaptation learning
- ✓ Achieve the ***State-of-the-Art performance*** compared to prior adaptation approaches and ***improve robustness of semantic segmentation models across views***

Cross-view Geometric Constraint



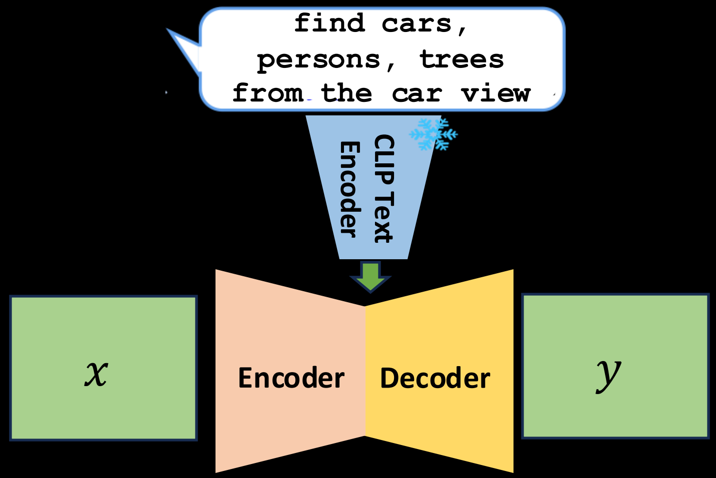
$$x_s = \mathcal{R}(K_s, [R_s, t_s], \Theta)$$

Cross-view Geometric Constraint



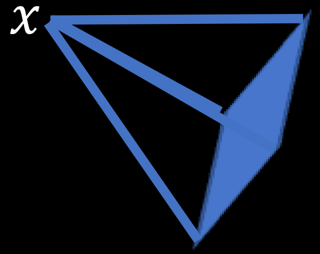
$$x_s = \mathcal{R}(K_s, [R_s, t_s], \Theta)$$

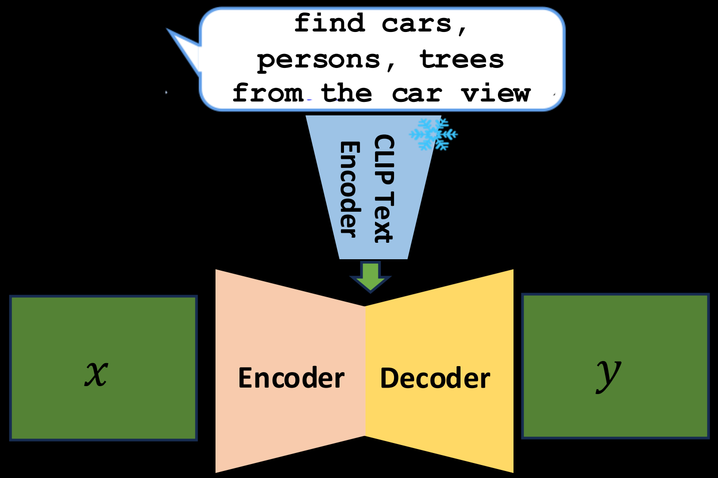
Cross-view Geometric Constraint



$x_s = \mathcal{R}(K_s, [R_s, t_s], \Theta)$

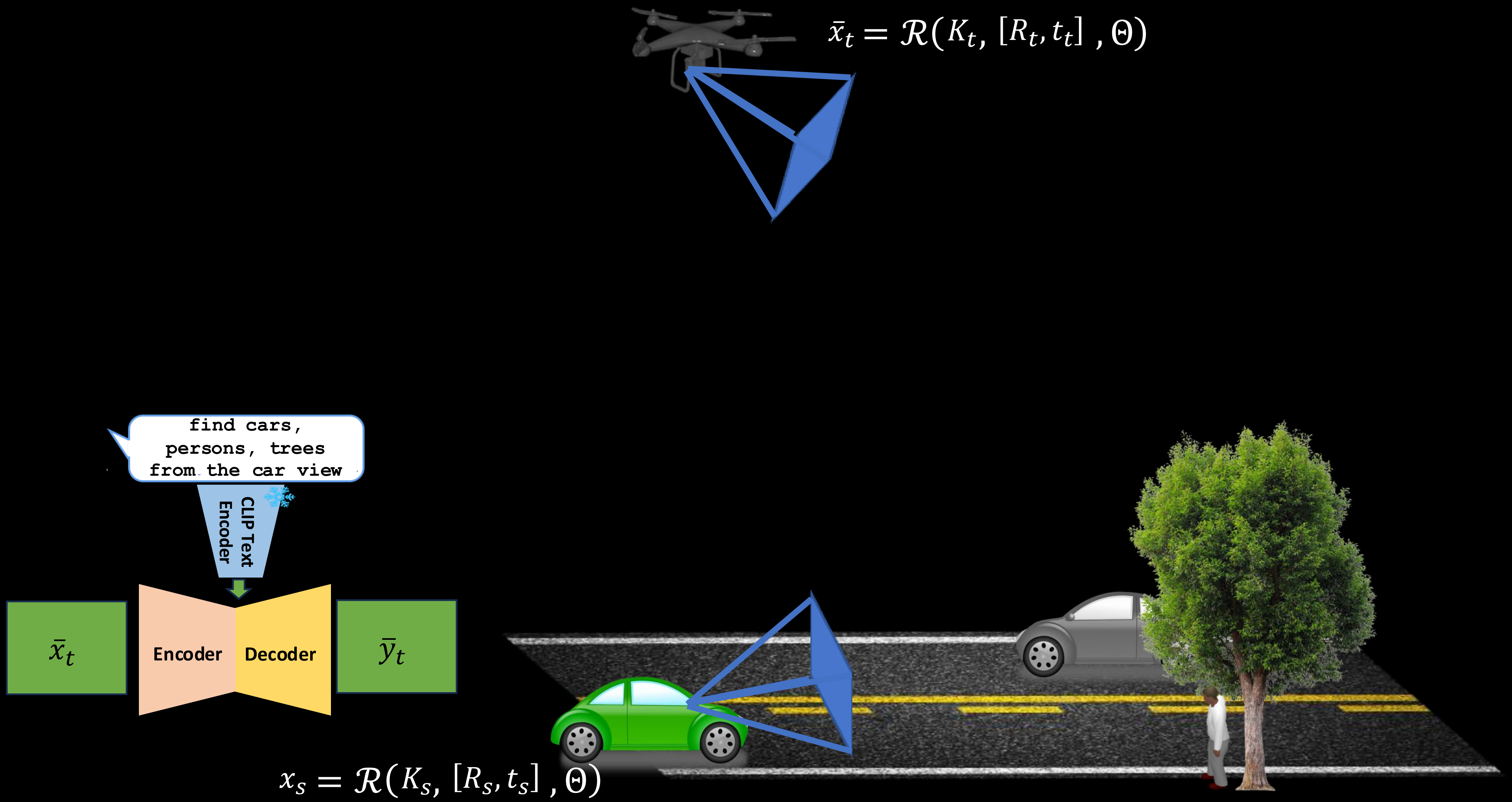
Cross-view Geometric Constraint





$$x_s = \mathcal{R}(K_s, [R_s, t_s], \Theta)$$

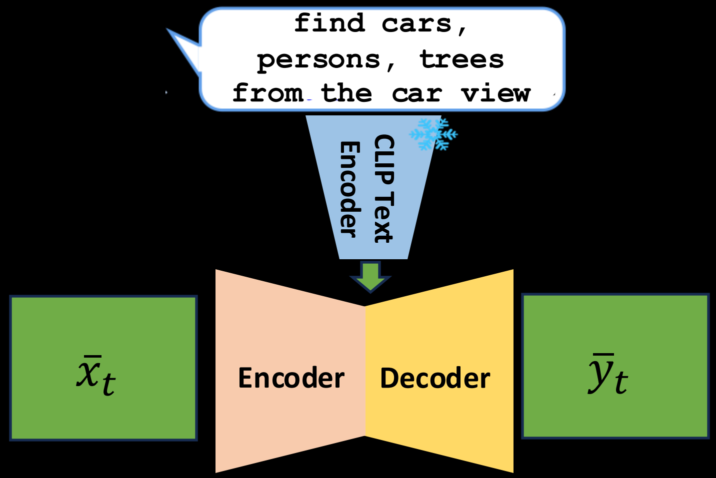
Cross-view Geometric Constraint



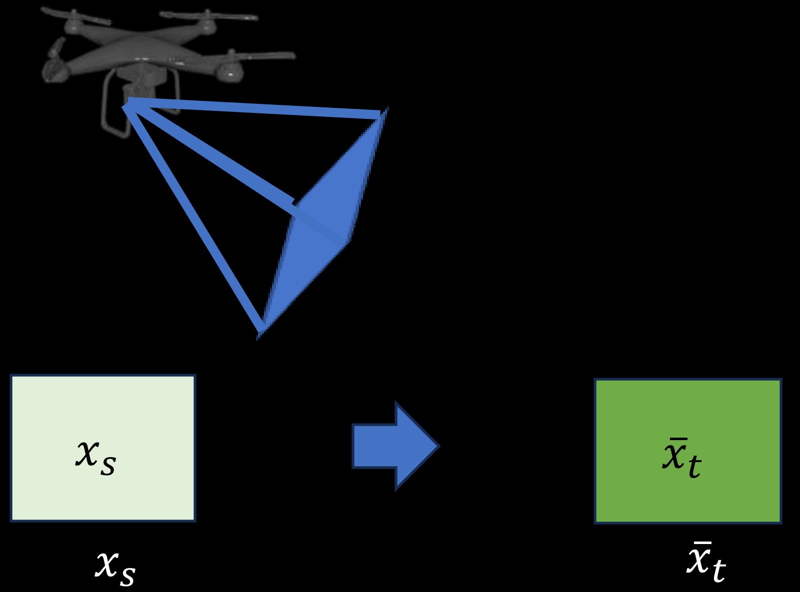
Cross-view Geometric Constraint

x_s $\bar{x}_t$

$$x_s = \mathcal{R}(K_s, [R_s, t_s], \Theta) \qquad \bar{x}_t = \mathcal{R}(K_t, [R_t, t_t], \Theta)$$

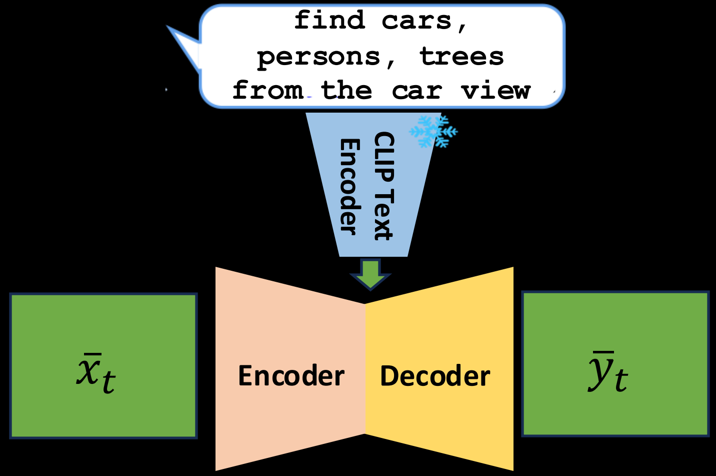


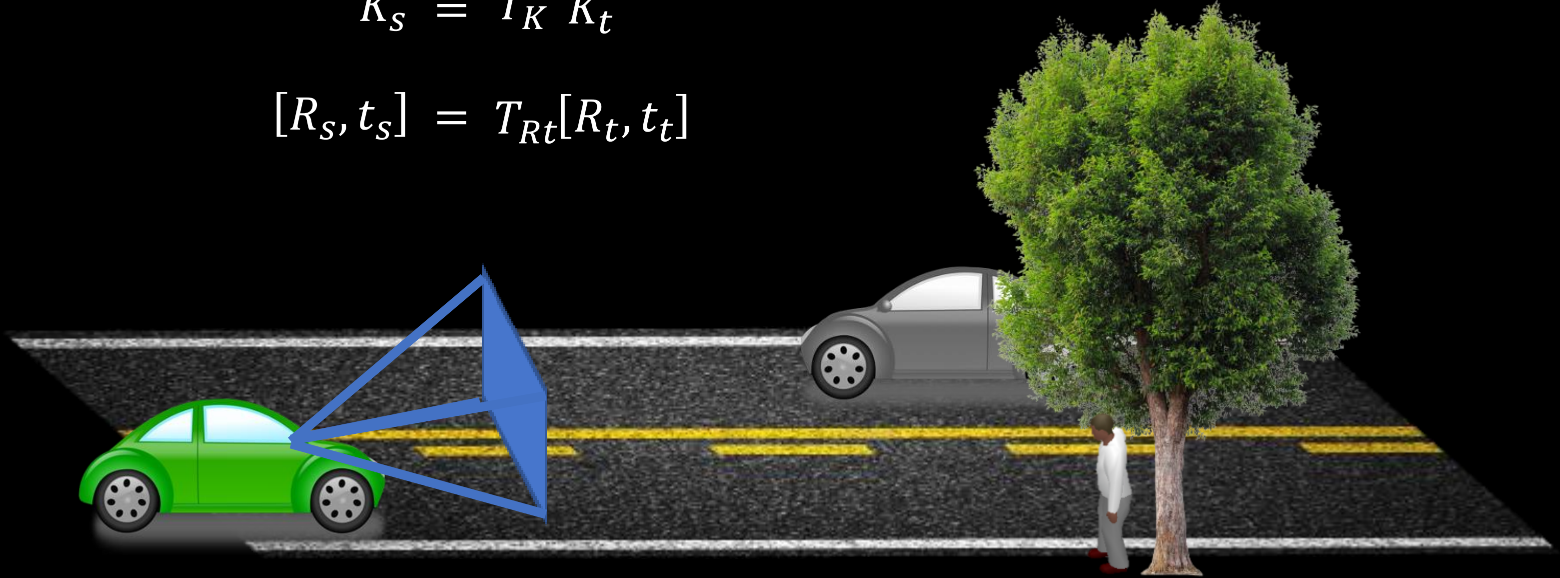
Cross-view Geometric Constraint



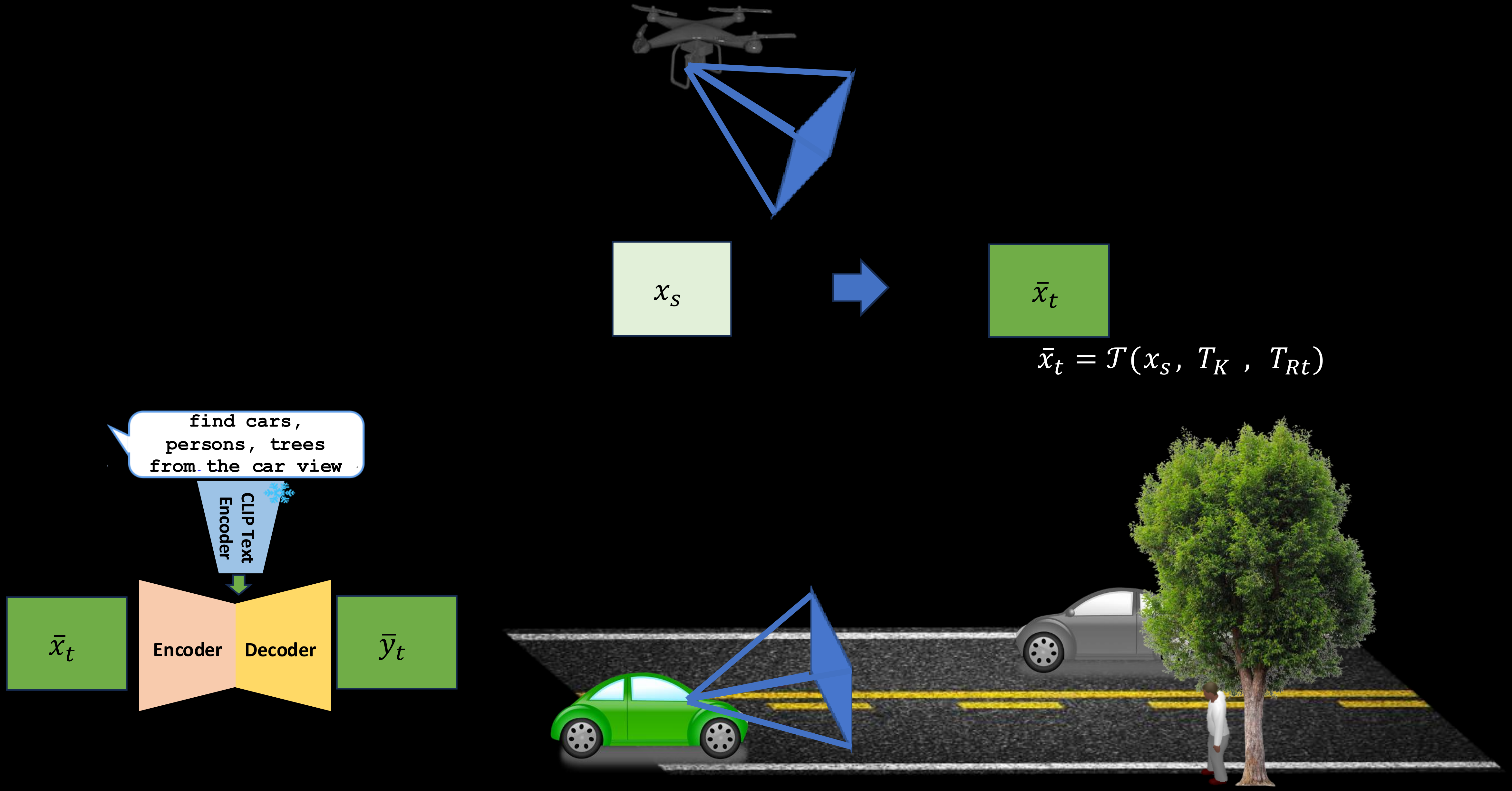
$$K_s = T_K \; K_t$$

$$[R_s, t_s] = T_{Rt}[R_t, t_t]$$

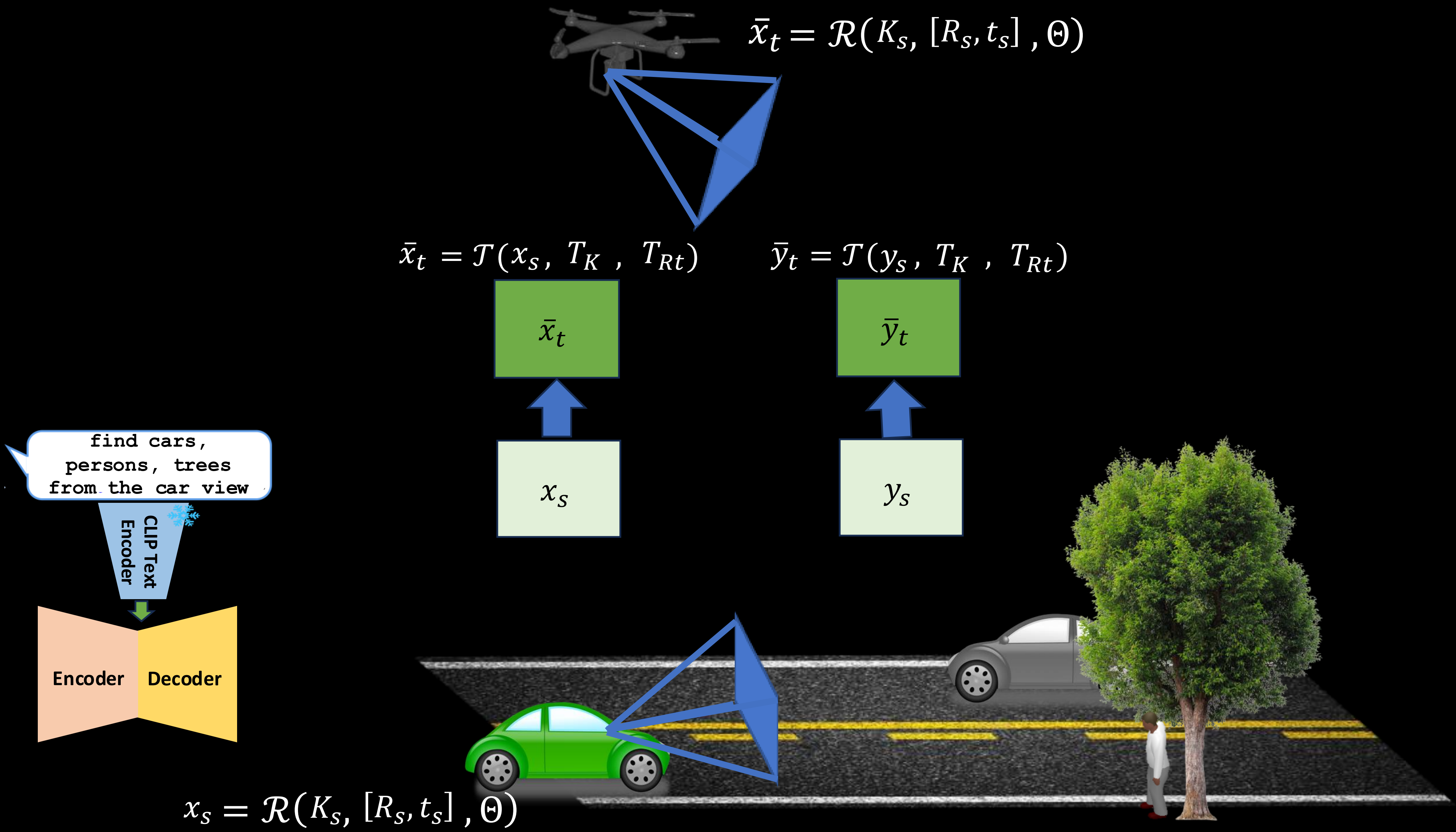




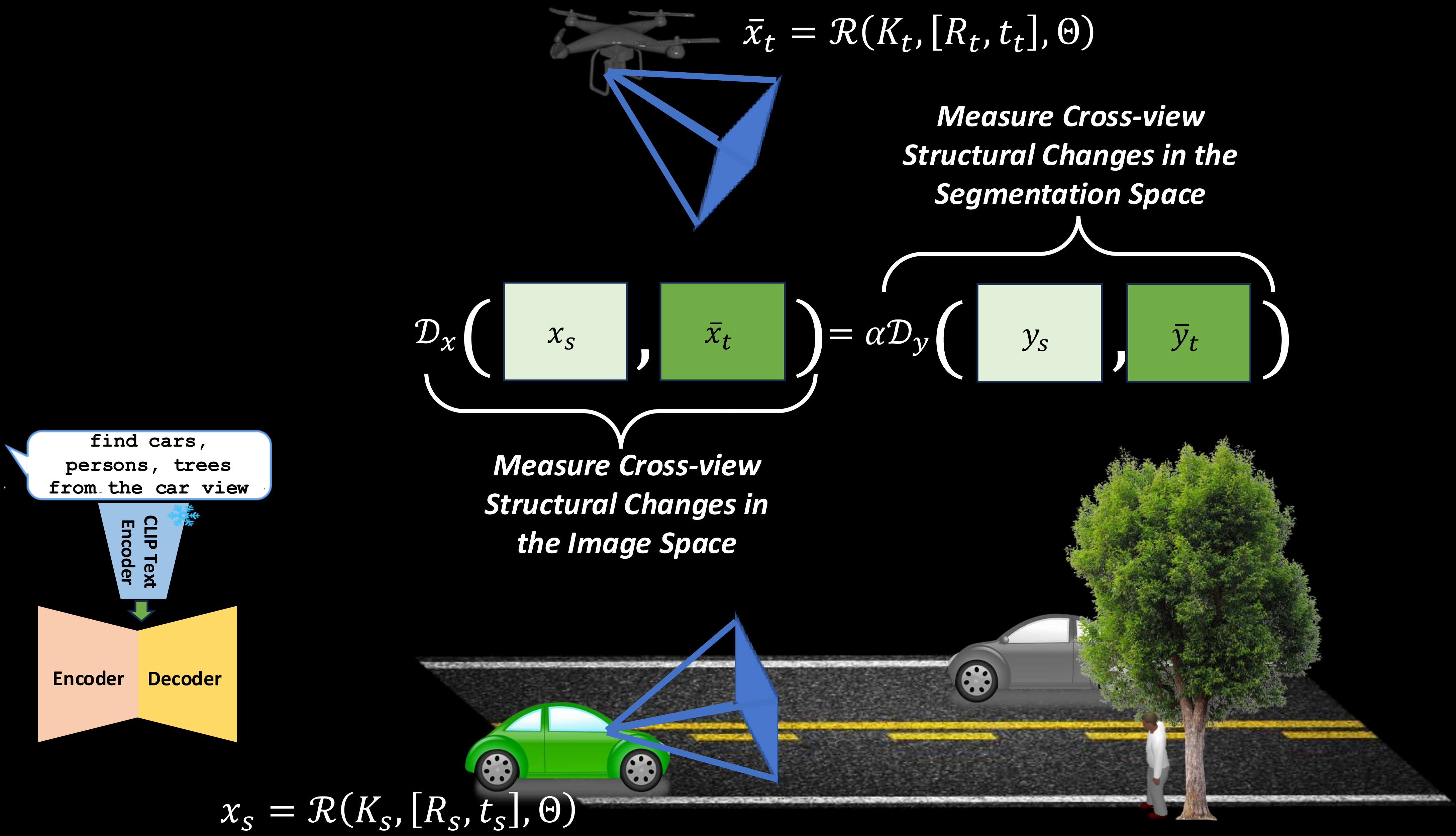
Cross-view Geometric Constraint



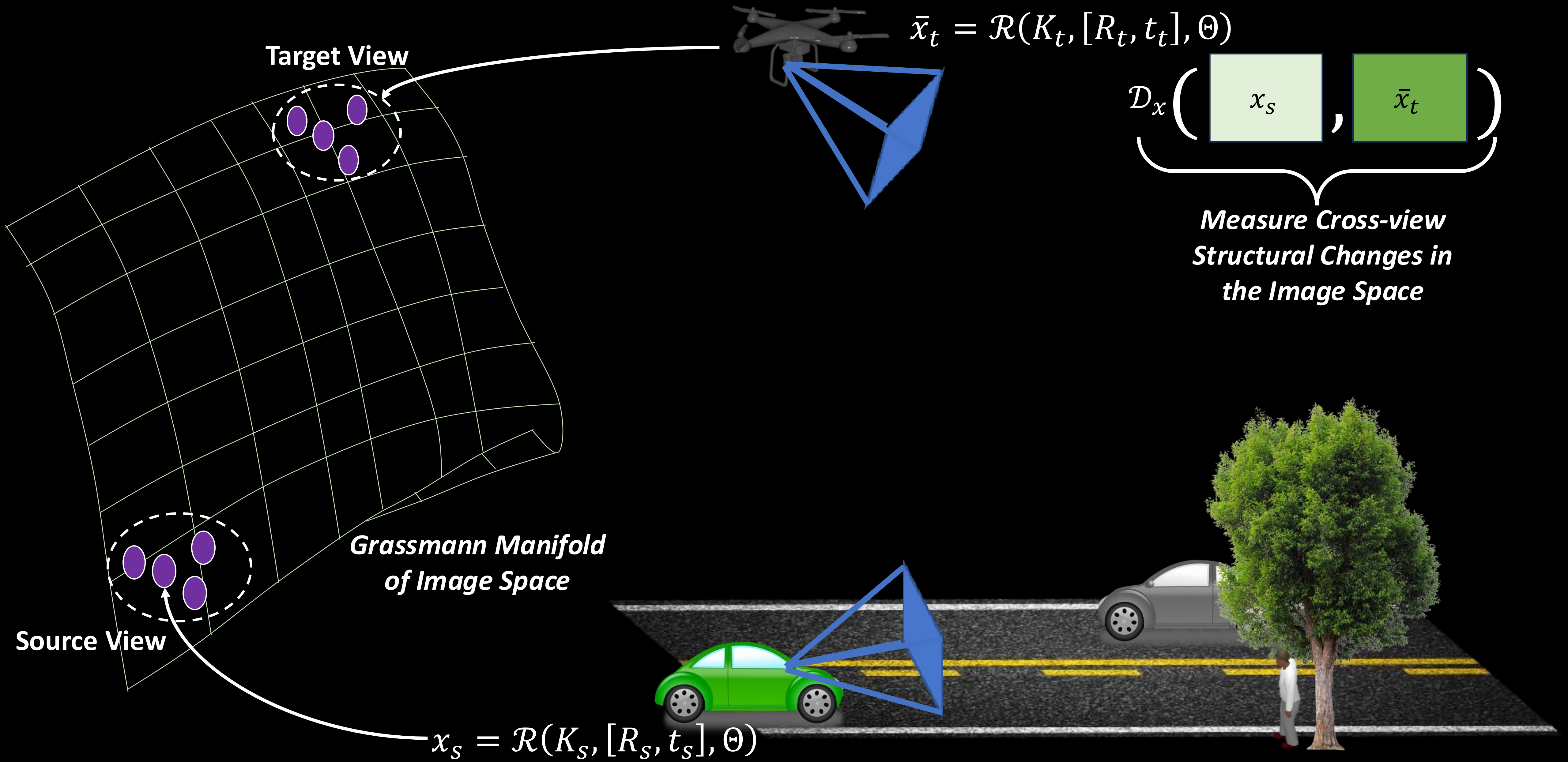
Cross-view Geometric Constraint



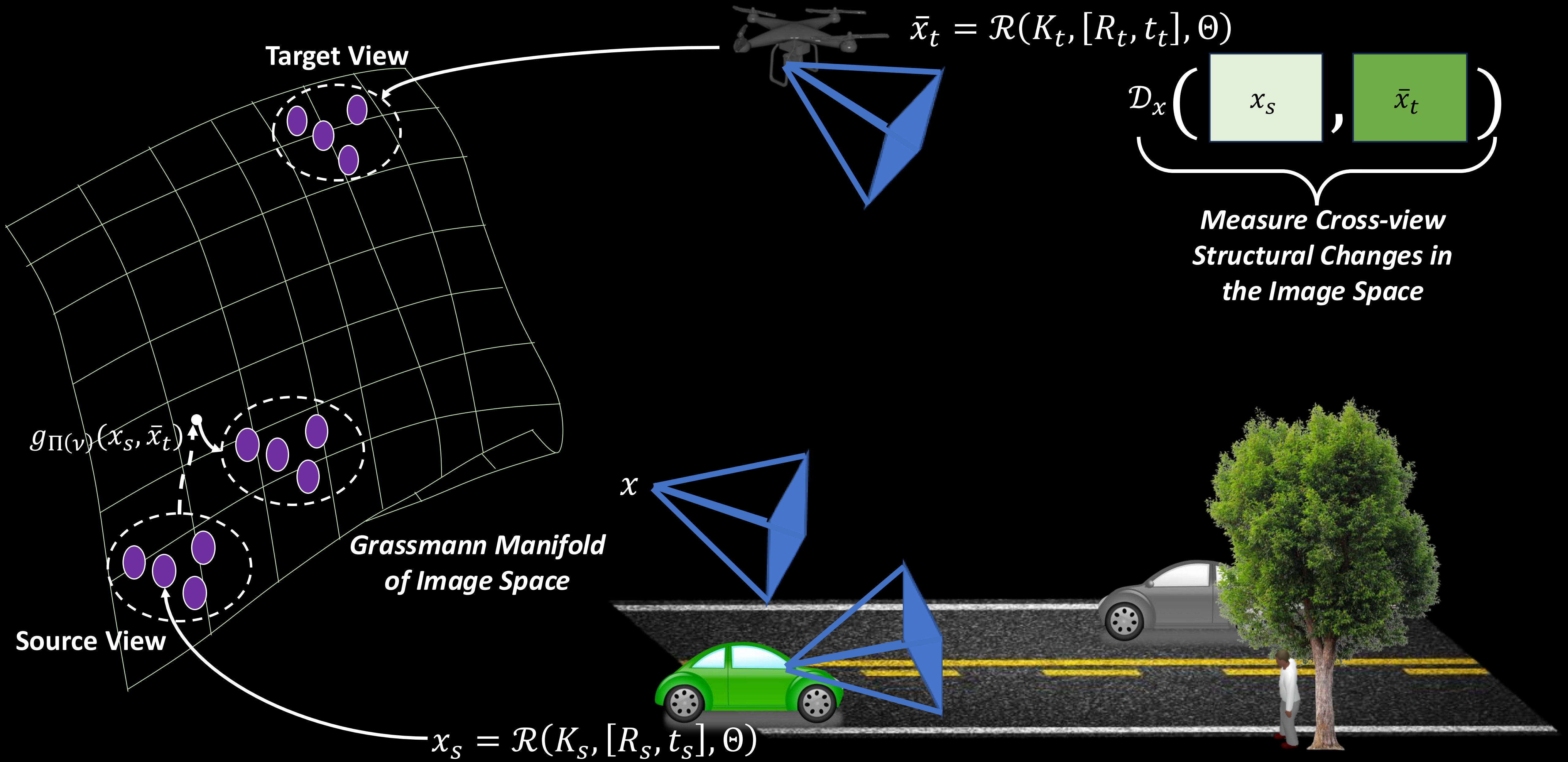
Cross-view Geometric Constraint



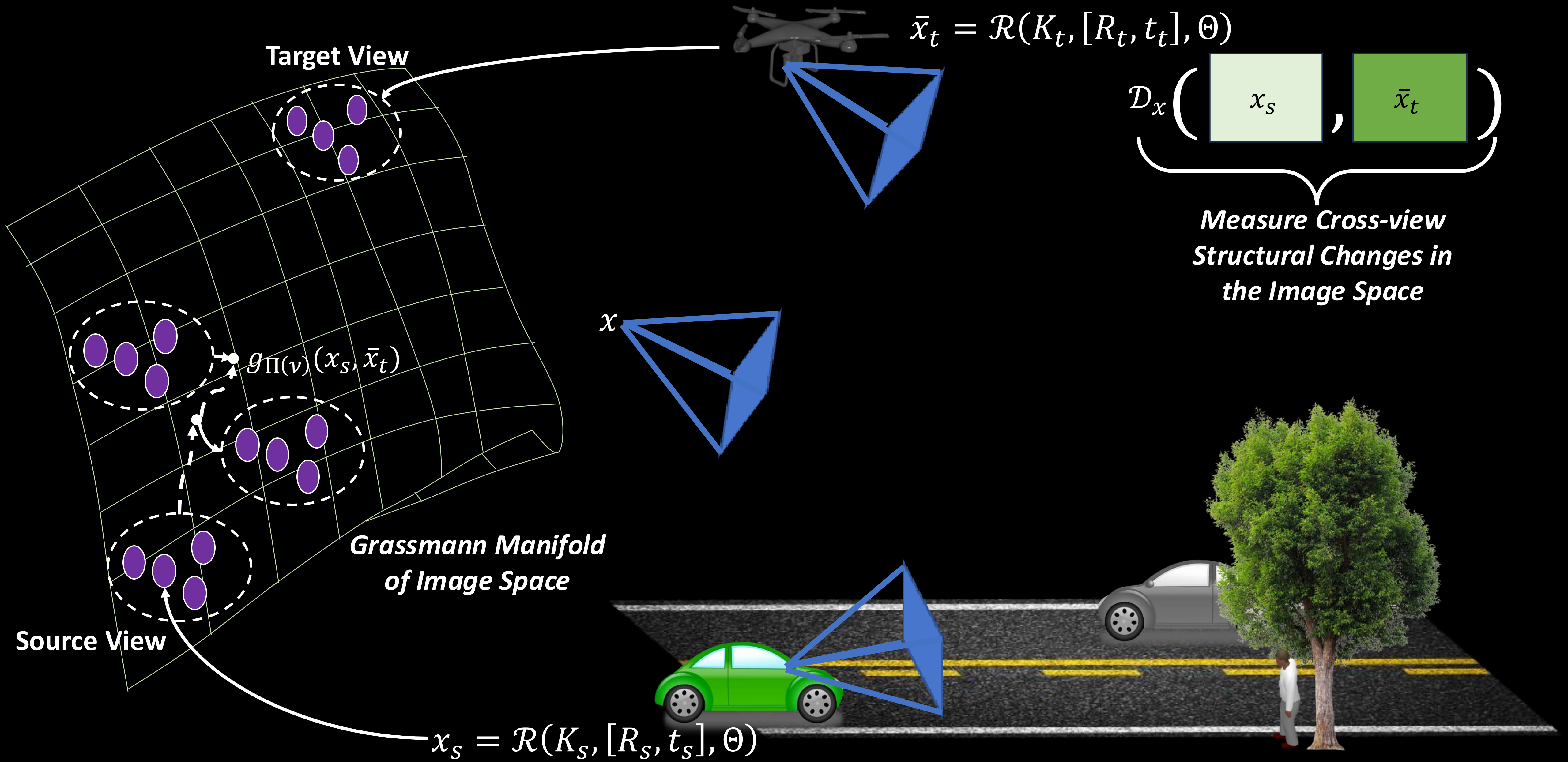
Cross-view Geometric Constraint



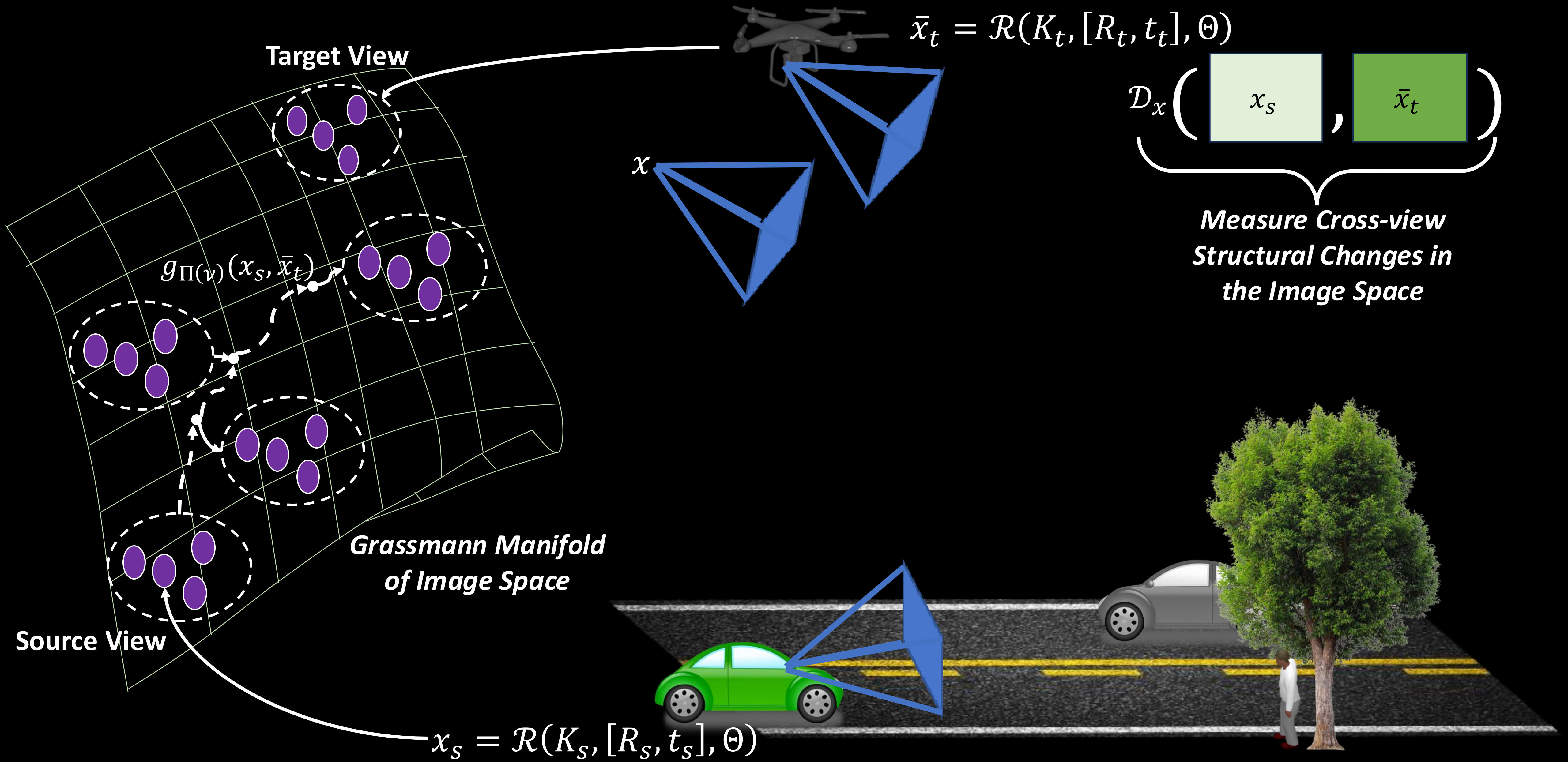
Cross-view Geometric Constraint



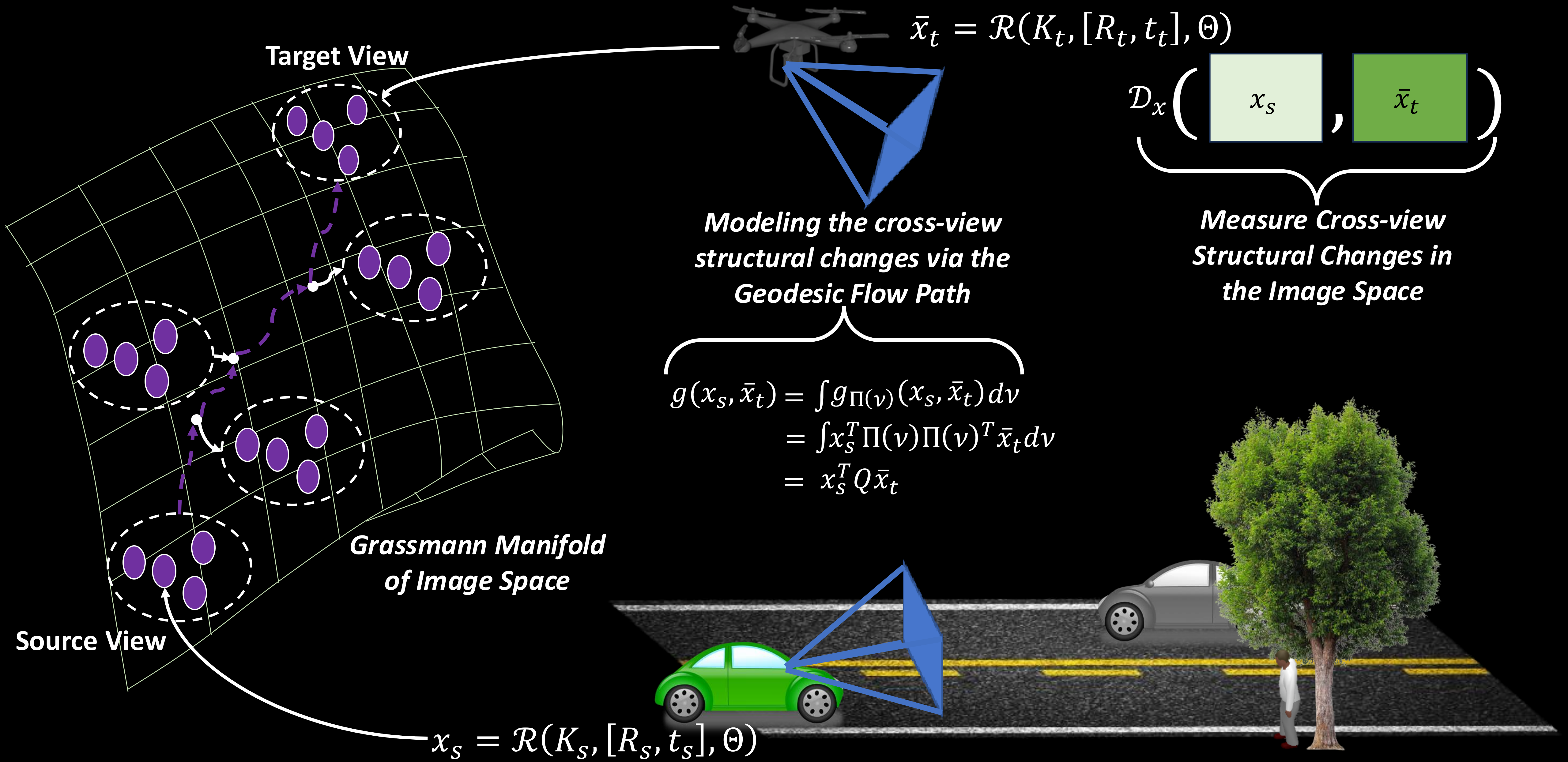
Cross-view Geometric Constraint



Cross-view Geometric Constraint

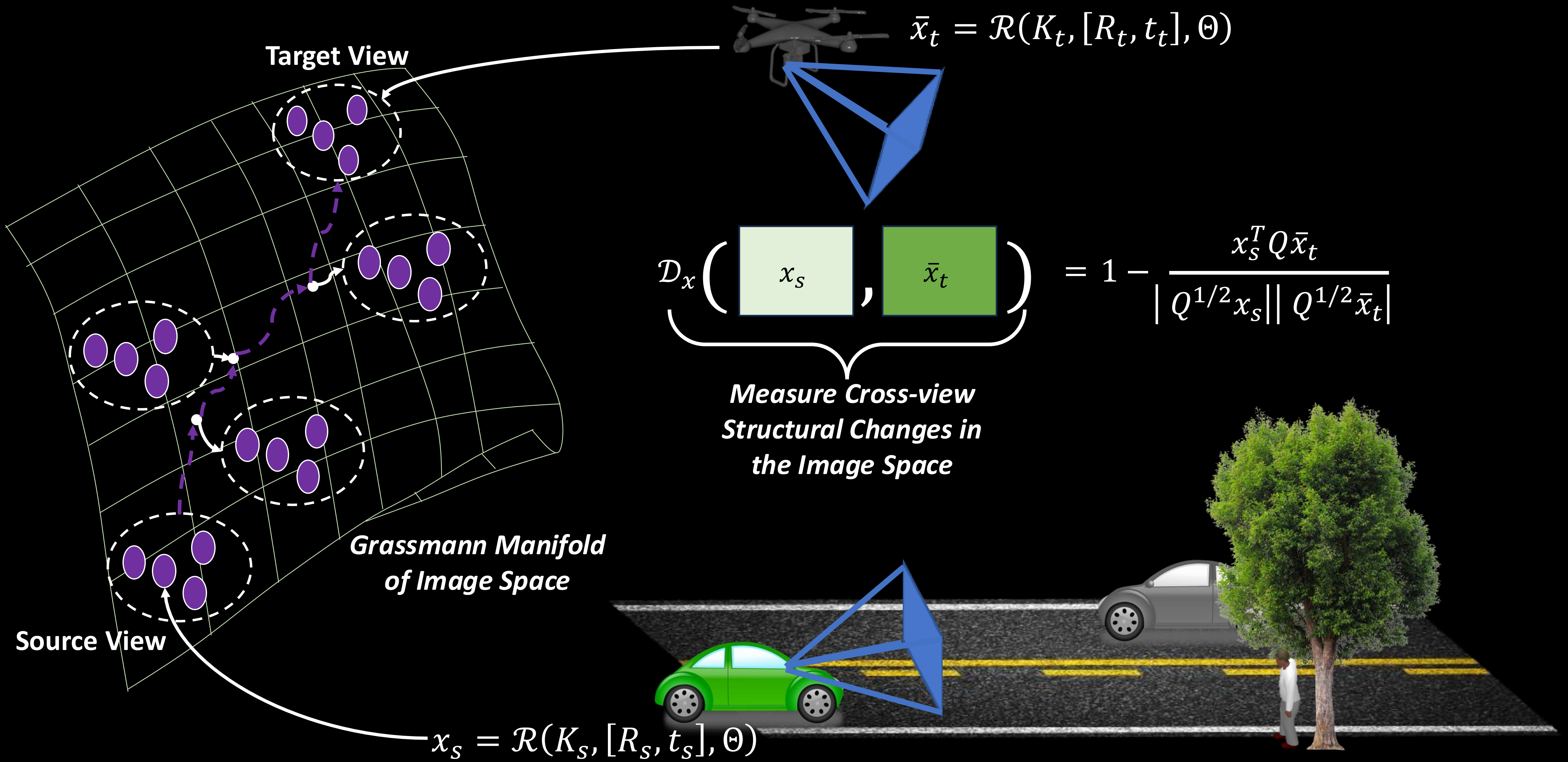


Cross-view Geometric Constraint



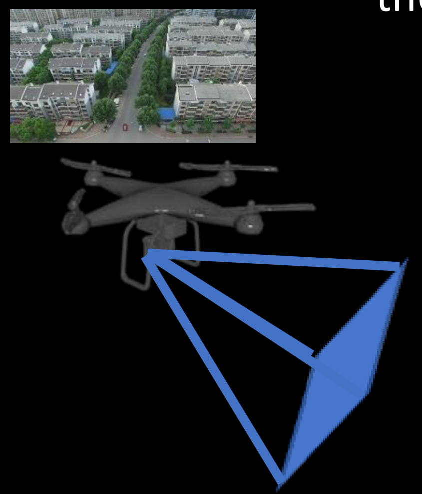
$$
\begin{aligned}
g(x_s, \bar{x}_t) &= \int g_{\Pi(\nu)}(x_s, \bar{x}_t)\, d\nu \\
&= \int x_s^T \Pi(\nu)\Pi(\nu)^T \bar{x}_t d\nu \\
&= x_s^T Q \bar{x}_t
\end{aligned}
$$

Cross-view Geometric Constraint



Cross-view Geometric Constraint

However, in practice, images of source and target views are collected independently (unpaired data)

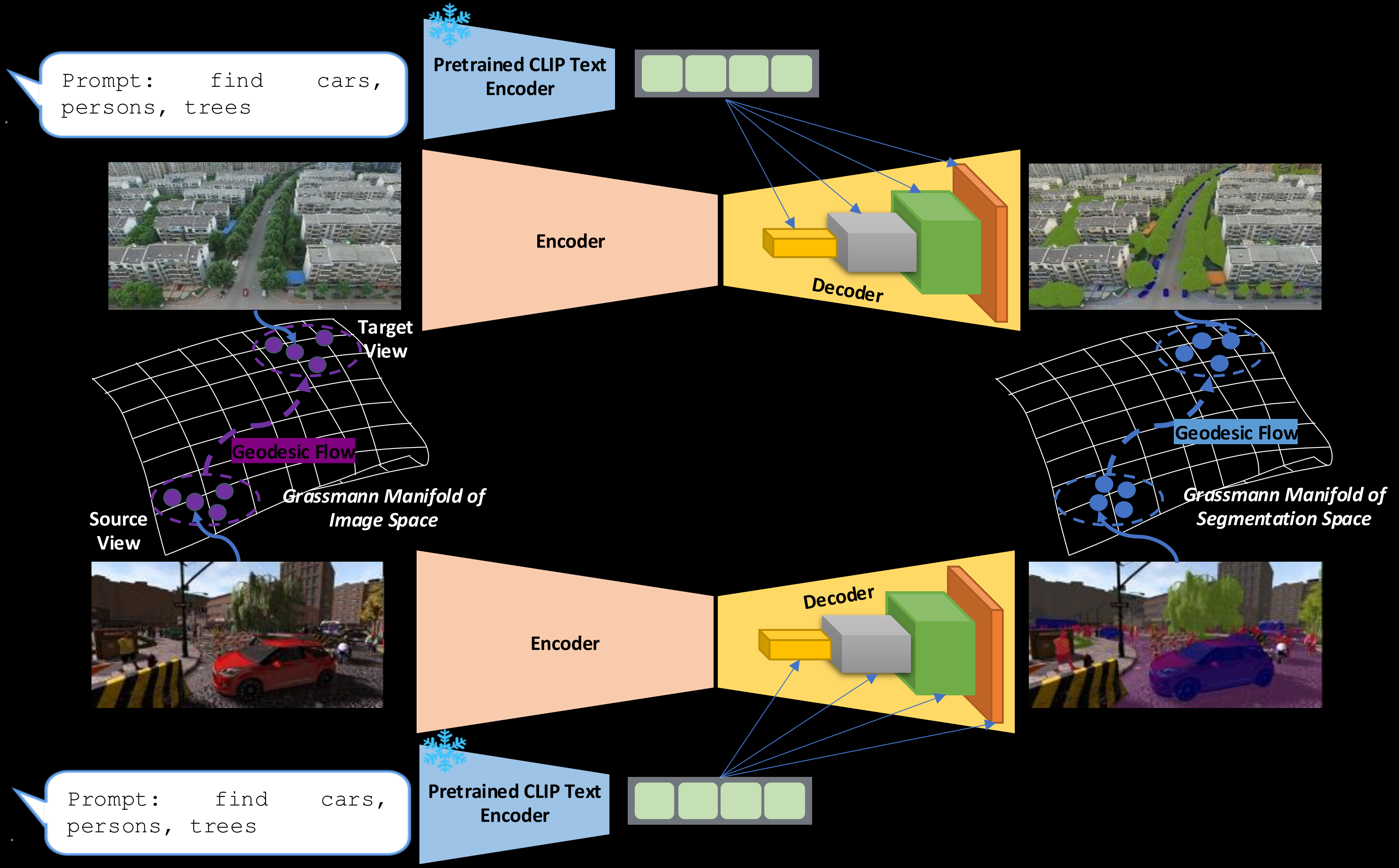


Cross-view Geometric Constraint

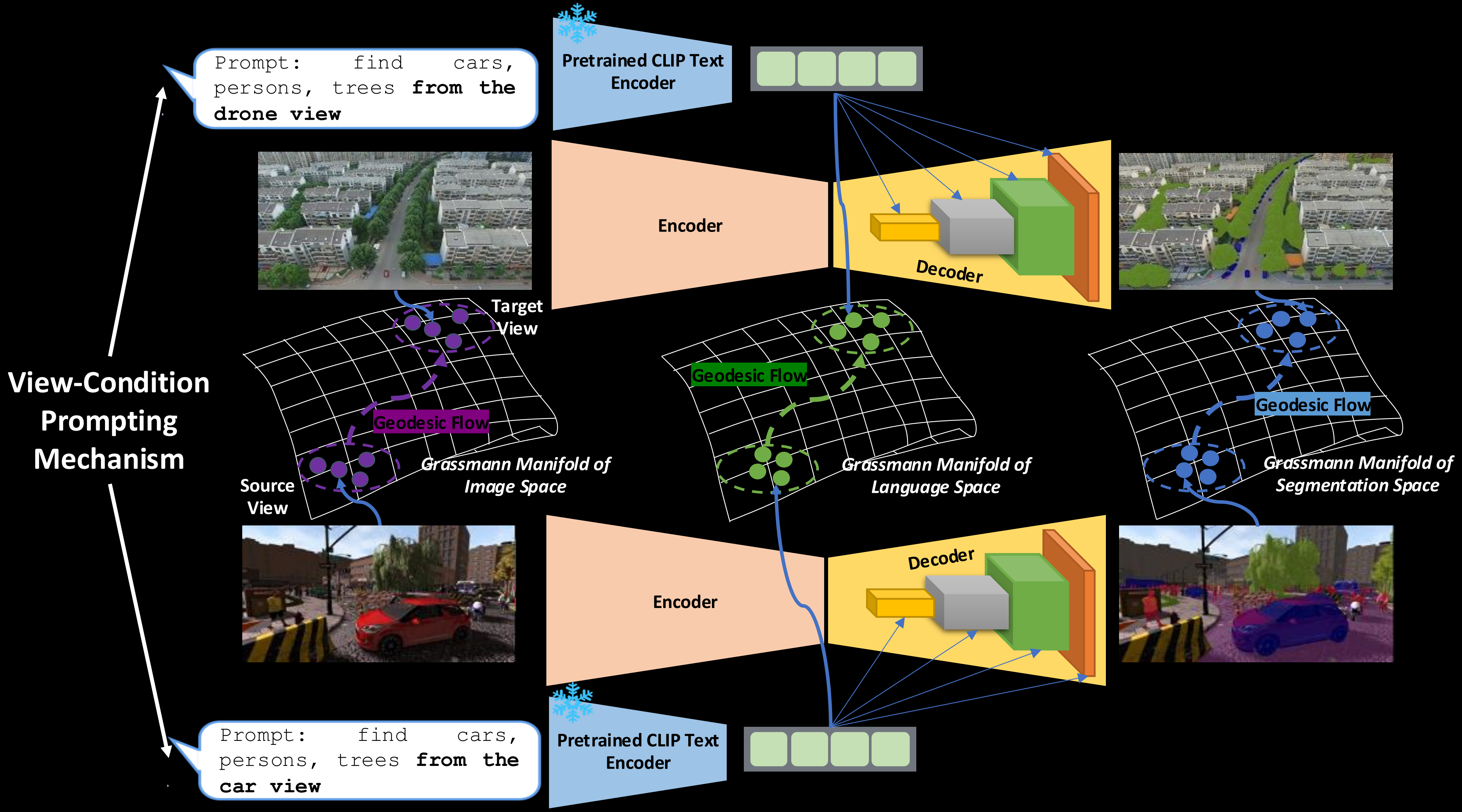
$\mathcal{D}_x\big($  ,  $\big) = \alpha\mathcal{D}_y\big($  ,  $\big)$

Modeling the Cross-view Geometric Constraint on Unpaired Data

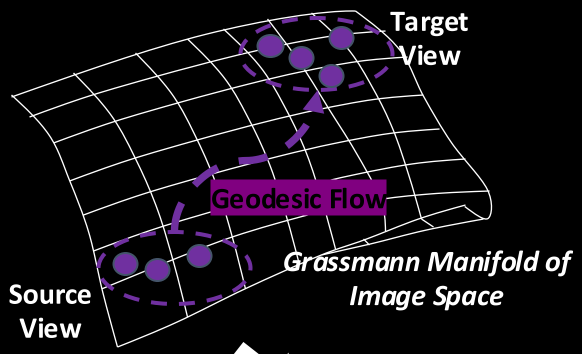
Cross-view Geometric Constraint

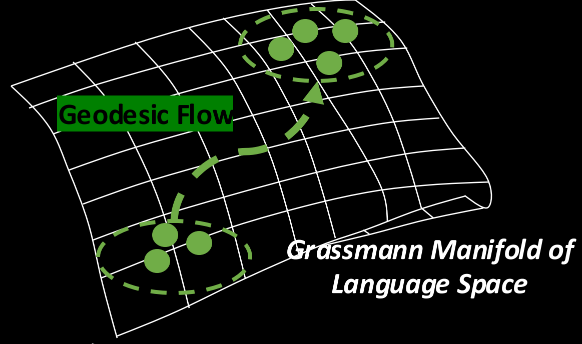


Cross-view Geometric Constraint



Cross-view Geometric Constraint



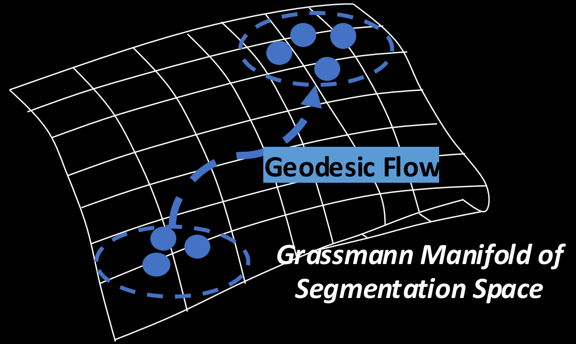


$$|\mathcal{D}_x(x_s, x_t) - \alpha \mathcal{D}_y(y_s, y_t)|$$

$$|\mathcal{D}_p(f_s^p, f_t^p) - \gamma \mathcal{D}_y(y_s, y_t)|$$

Cross-view Geometric Adaptation Loss Guided by the Geodesic Flow in the Image Space

Cross-view Geometric Adaptation Loss Guided by the Geodesic Flow in the Language Space



Video Demo