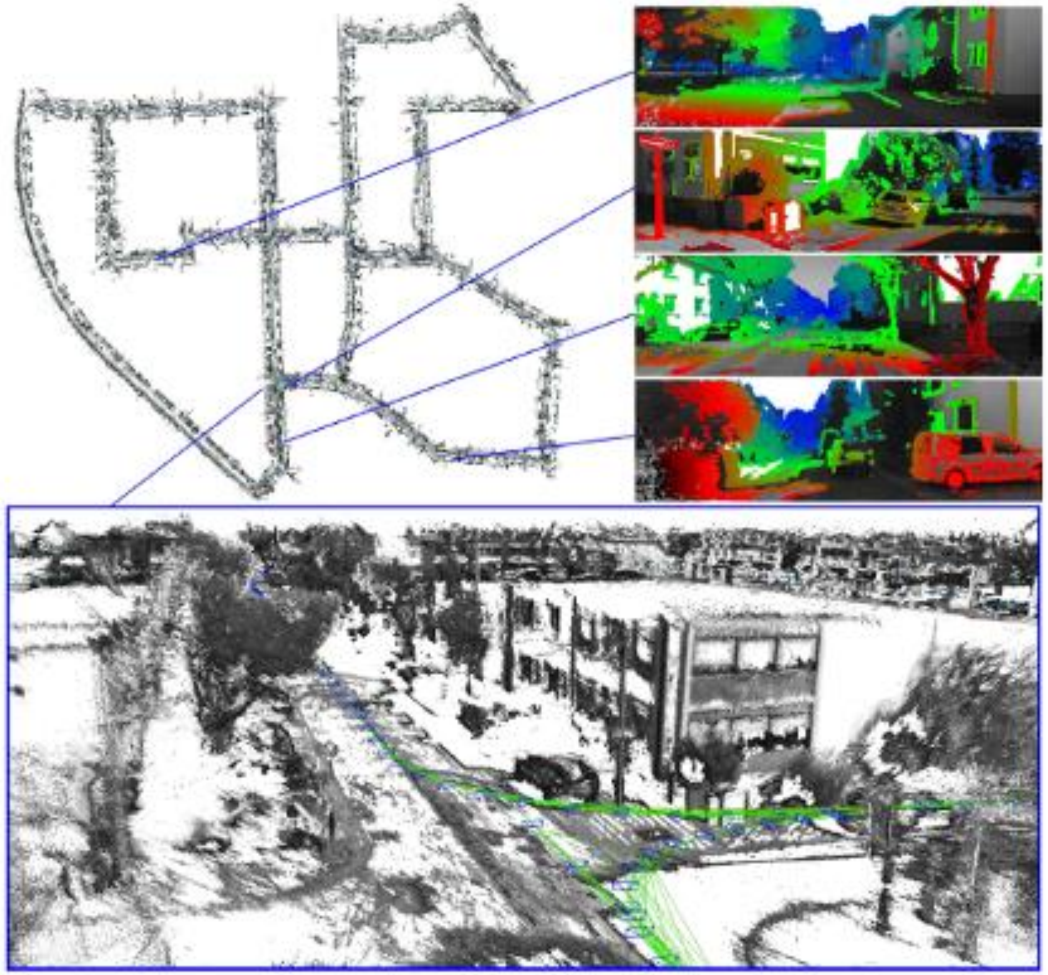
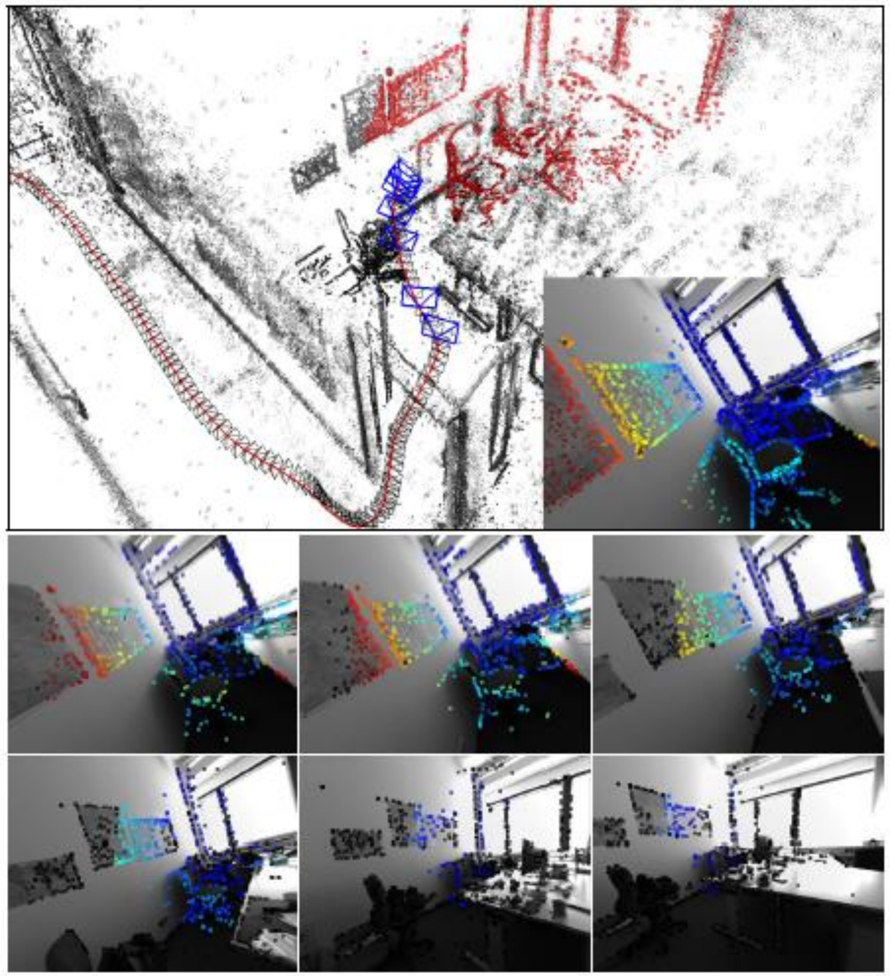
PHOTOMETRIC ODOMETRY FOR DYNAMIC OBJECTS

Anton Troynikov

MOTIVATION: LARGE COMPLEX SENSOR RIGS

INSPIRATION: DIRECT PHOTOMETRIC SLAM



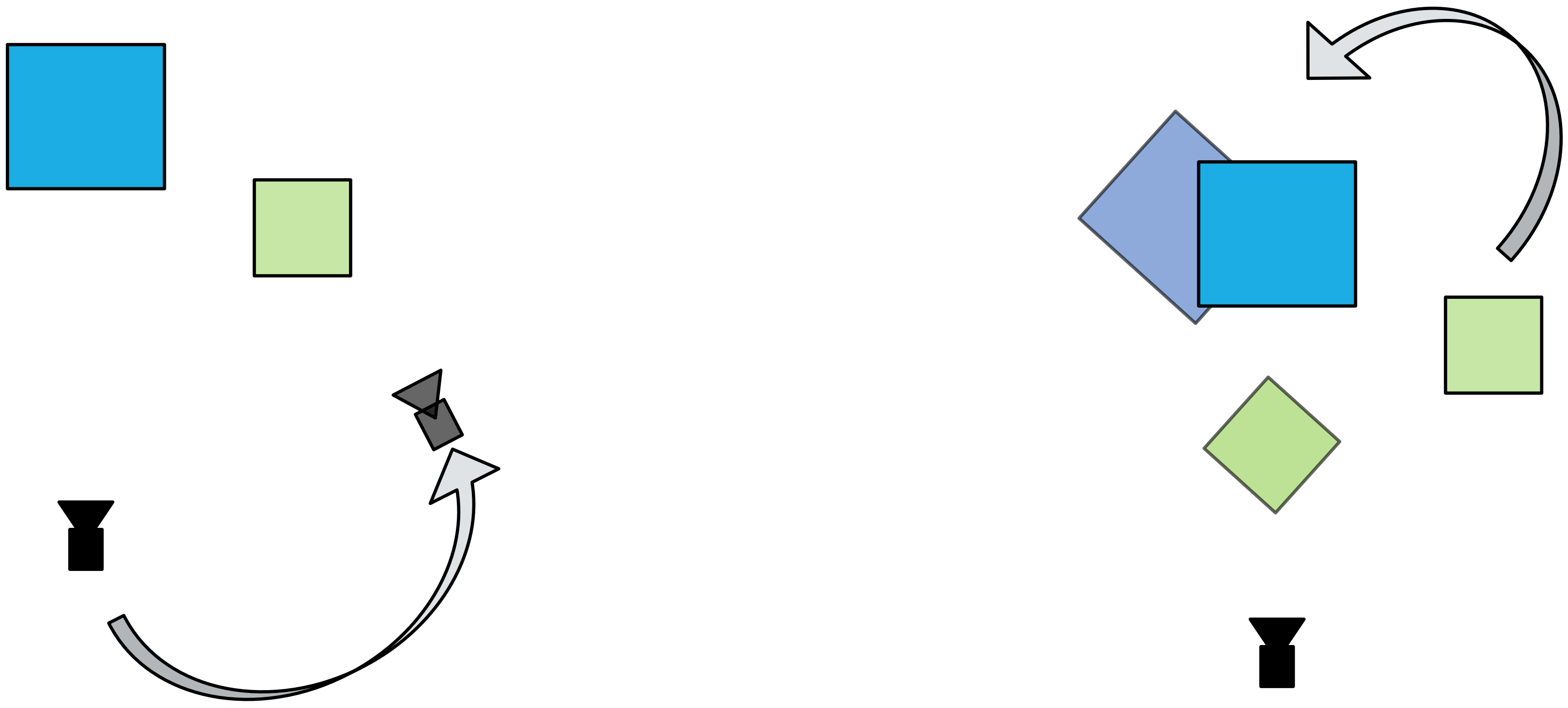


PROBLEM STATEMENT

Is it possible to:

- Use data from visible light cameras
- Estimate the motion of dynamic objects in the world
- Estimate the motion of the camera itself

KEY OBSERVATION: TRANSFORM EQUIVALENCE



Observing of a static object from a moving camera will be the same as a moving object observed from a static camera.

DIRECT PHOTOMETRIC ALIGNMENT

$$r\left(\boldsymbol{x}_{\mathbf{1}}, d_1, \boldsymbol{T}\right) := \boldsymbol{I}_2\left(\tau\left(\boldsymbol{x}_1, d_1, \boldsymbol{T}\right)\right) - \boldsymbol{I}_1\left(\boldsymbol{x}_1\right)$$

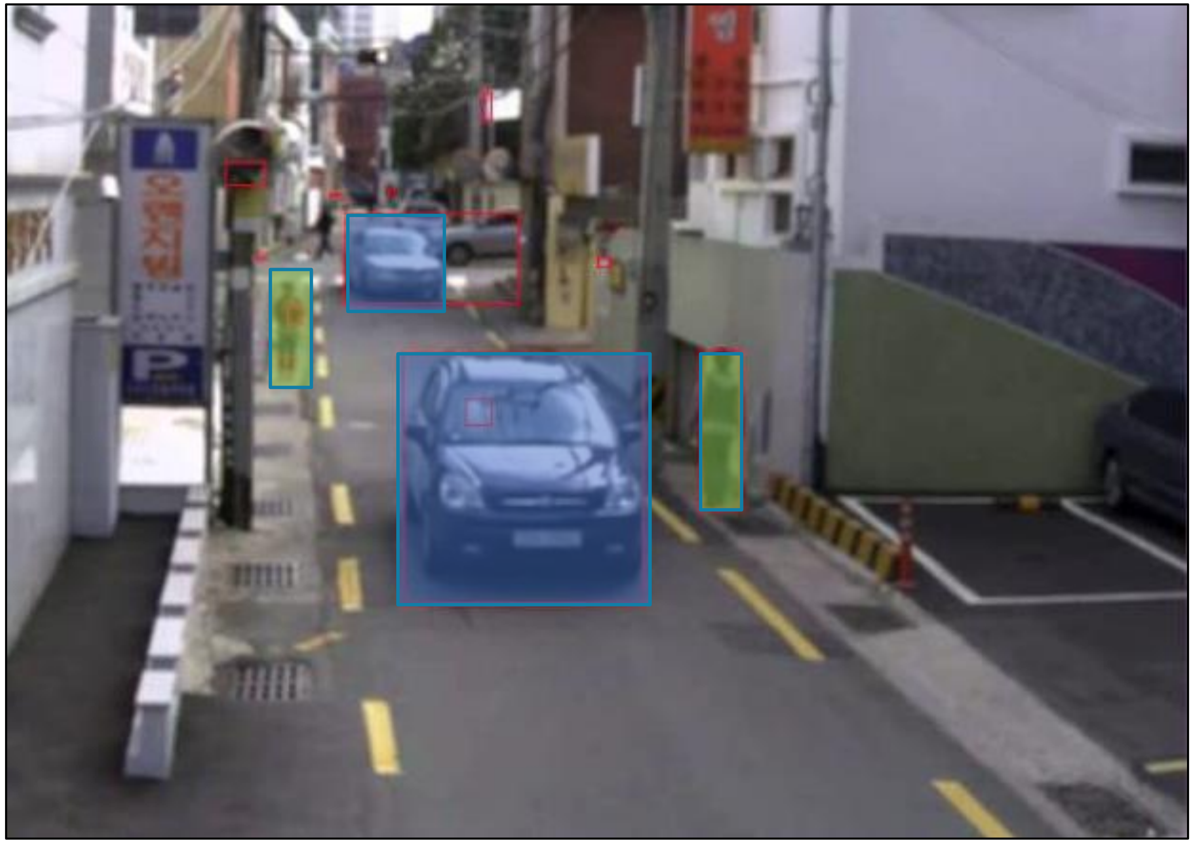
$$\boldsymbol{T}_{MAP} = \arg\max_{\boldsymbol{T}} p\left(\boldsymbol{T} \middle| \boldsymbol{r}\right)$$

$$\boldsymbol{T}_{MAP} = \arg\min_{\boldsymbol{T}} \sum_i w\left(r_i\right)\left(r_i\left(\boldsymbol{T}\right)\right)^2$$

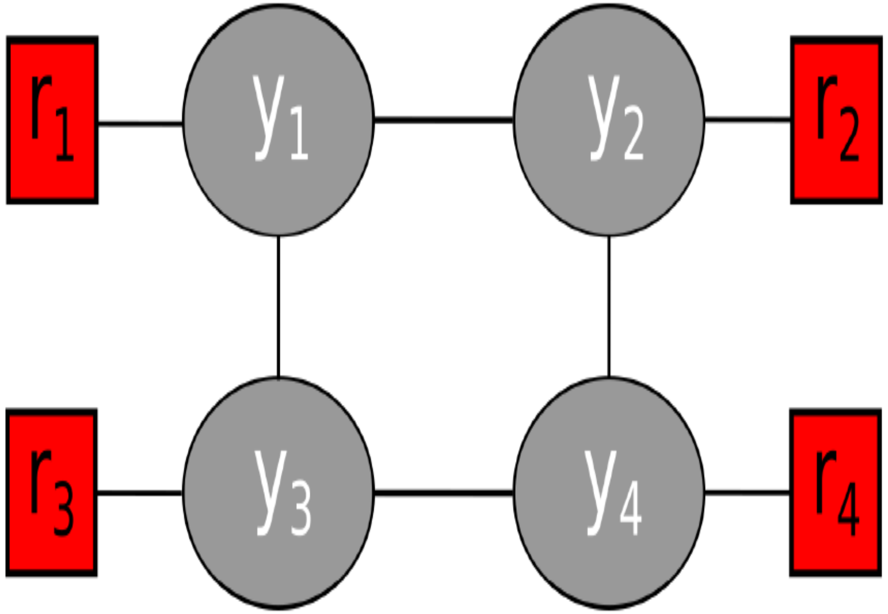
$$\hat{\boldsymbol{\xi}}_{MAP} = \arg\min_{\hat{\boldsymbol{\xi}}} \sum_i w\left(r_i\right)\left(r_i\left(\hat{\boldsymbol{\xi}}\right)\right)^2$$

KEY OBSERVATION: MOTION OUTLIERS

Independently moving objects produce large photometric residuals between image frames.



MOTION SEGMENTATION – RANDOM FIELDS



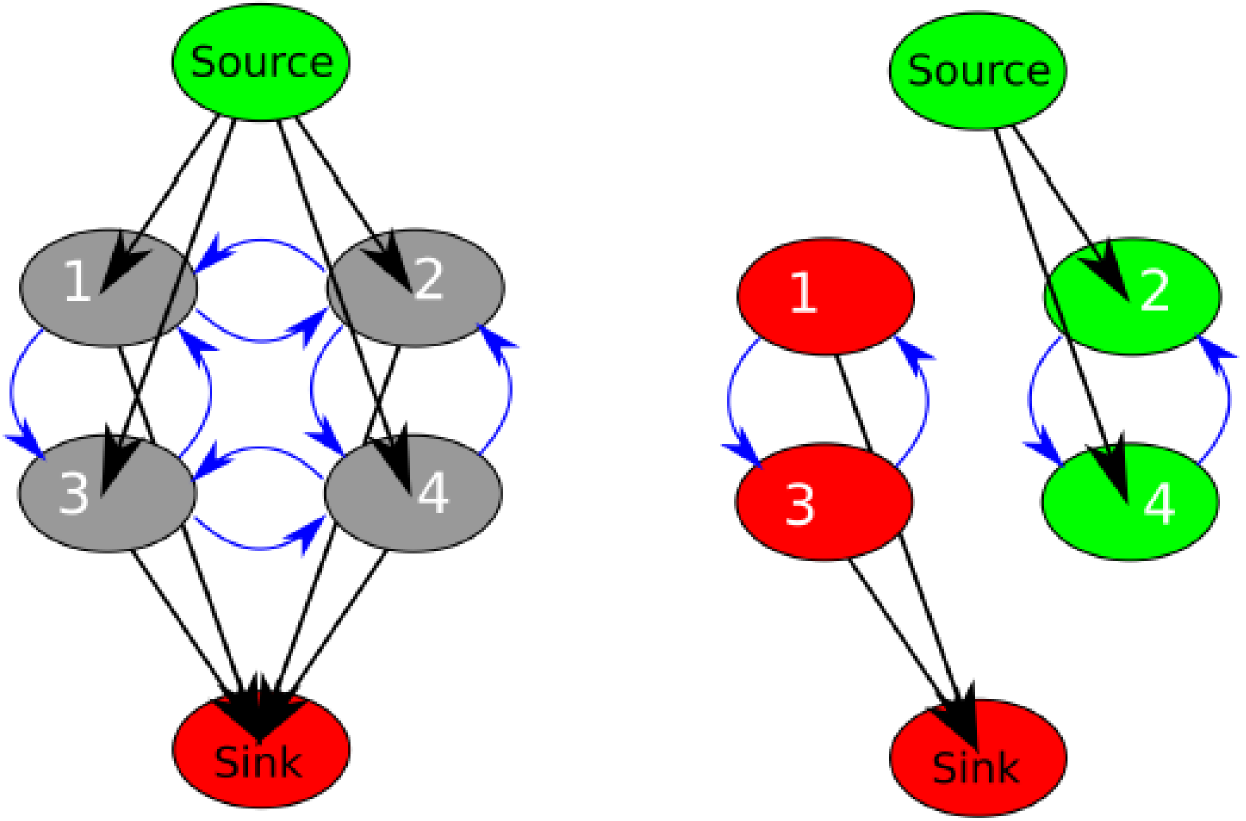
$$p(\boldsymbol{Y}|\boldsymbol{R}) = \frac{1}{Z(\boldsymbol{R})}\tilde{p}(\boldsymbol{Y}|\boldsymbol{R})$$

$$\tilde{p}(\boldsymbol{Y}|\boldsymbol{R}) = \prod_{i}^{n\times m} \phi(y_i, r_i) \prod_{j\in \mathrm{N}(\boldsymbol{x}_i)} \phi(y_i, y_j)$$

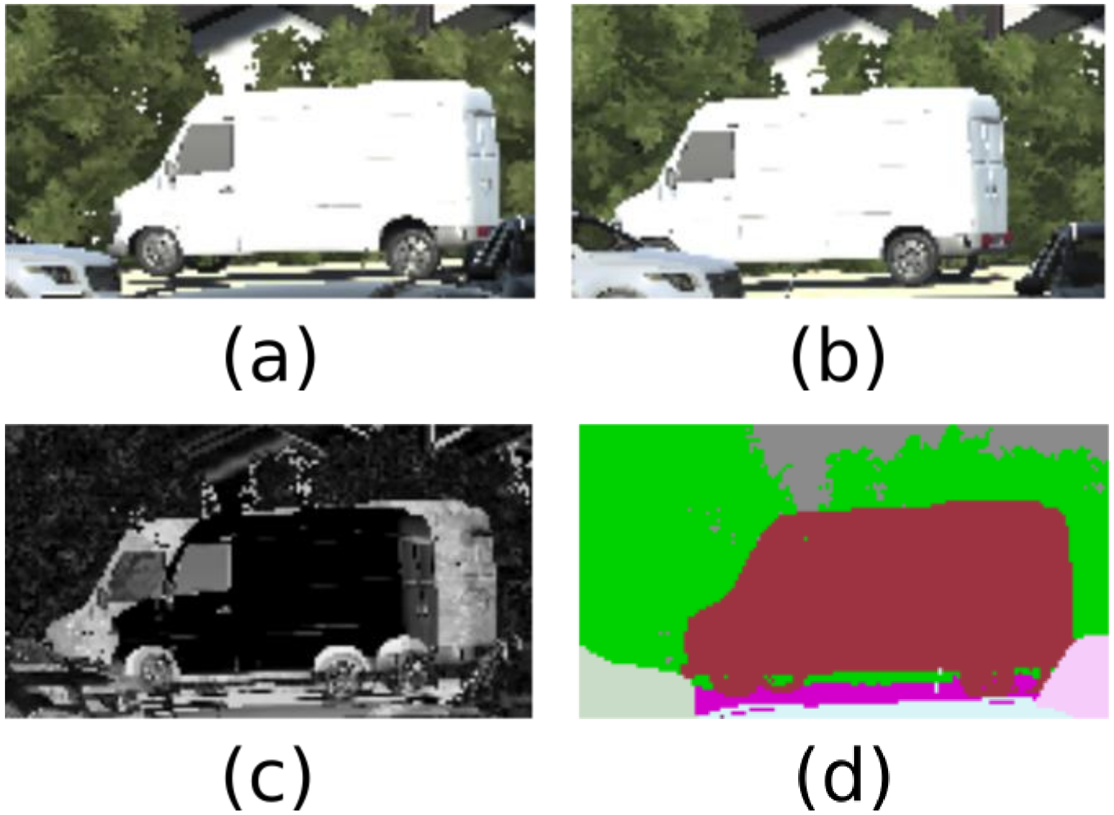
$$Z(\boldsymbol{R}) = \sum_{\boldsymbol{Y}\in\mathcal{Y}} \tilde{p}(\boldsymbol{Y}|\boldsymbol{R})$$

Magnitude of photometric residual as unary energy term in MRF segmentation.

MOTION SEGMENTATION – GRAPH CUTS



DISAMBIGUATION

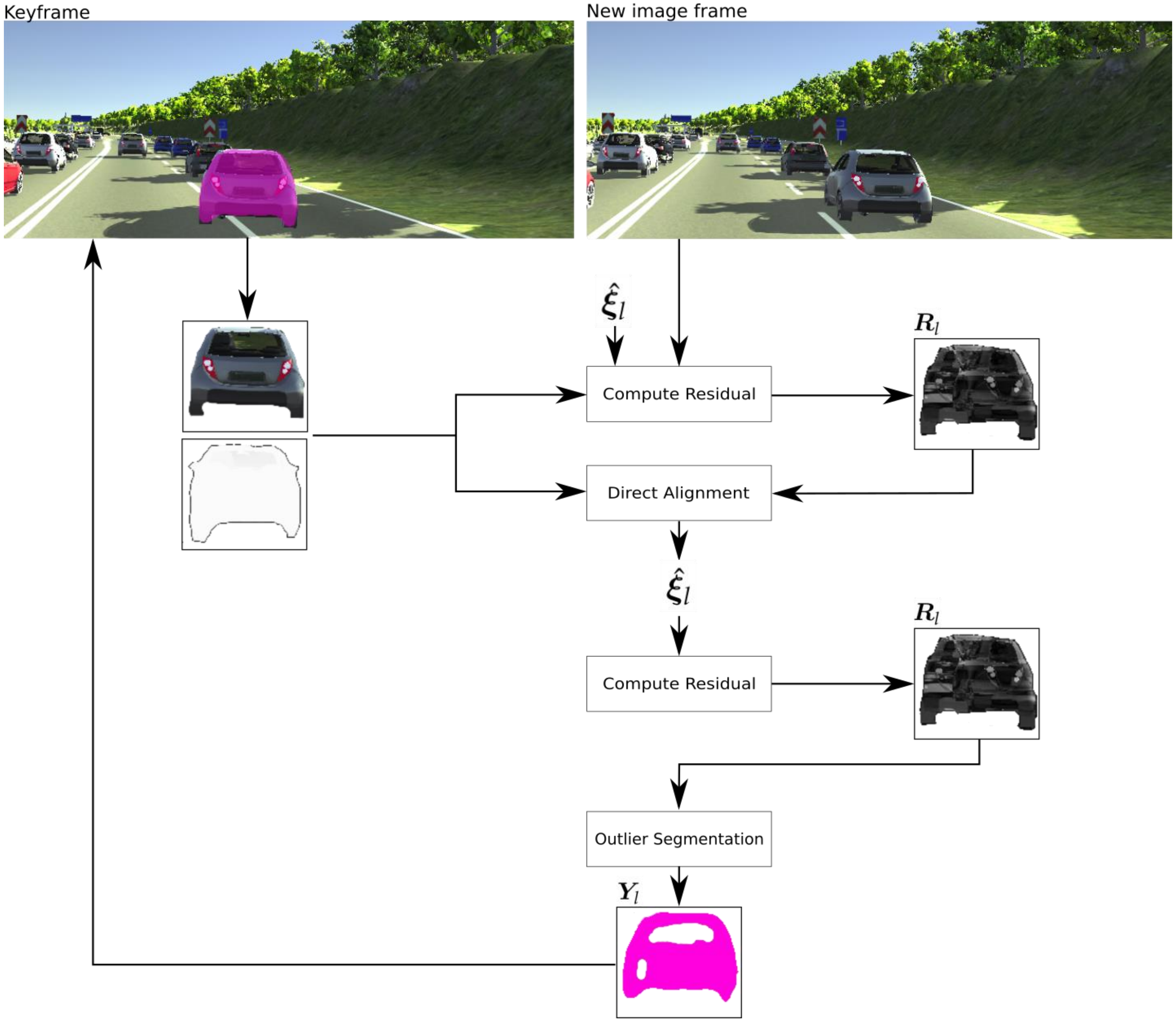


Instance segmentation to disambiguate object from background, including low-texture regions.

PUTTING IT TOGETHER: JOINT ESTIMATION

$$\xi_{lMAP} = \arg\max_{\hat{\xi}_l} p\left(\hat{\xi}_l | \boldsymbol{R}_l\right) = \arg\max_{\hat{\xi}_l} \sum_{\mathbf{Y}_l}^{\mathcal{Y}_l} p\left(\hat{\xi}_l, \mathbf{Y}_l | \boldsymbol{R}_l\right)$$

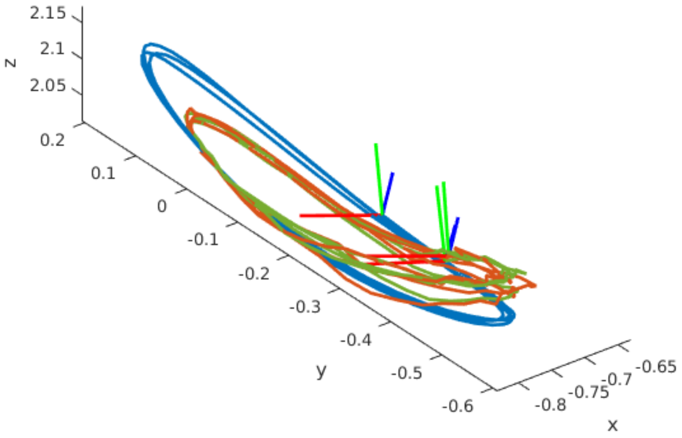
$\mathcal{I}_l \leftarrow \mathcal{X}_l$
$\boldsymbol{R}_l \leftarrow R(\boldsymbol{F}, \mathcal{X}_l, \hat{\xi}_l)$
while not converged **do**
$\quad \hat{\xi}_l \leftarrow \arg\max_{\hat{\xi}_l} p\left(\hat{\xi}_l | \boldsymbol{R}_l; \mathcal{I}_l\right)$
$\quad \boldsymbol{R}_l \leftarrow R(\boldsymbol{F}, \mathcal{X}_l, \hat{\xi}_l)$
$\quad \boldsymbol{Y}_l \leftarrow \arg\max_{\mathbf{Y}_l} p\left(\boldsymbol{Y}_l | \boldsymbol{R}_l\right)$
$\quad \mathcal{I}_l \leftarrow \mathrm{assign}\left(\boldsymbol{Y}_l\right)$
end while

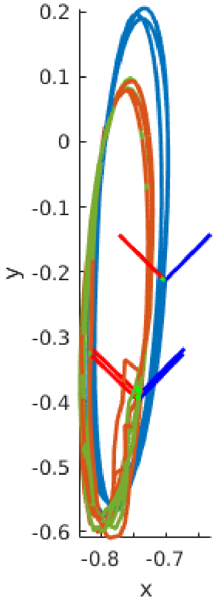


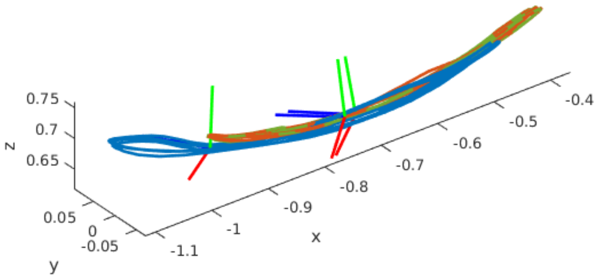
Keyframe
New image frame
$\hat{\xi}_l$
Compute Residual
R_l
Direct Alignment
$\hat{\xi}_l$
Compute Residual
R_l
Outlier Segmentation
Y_l

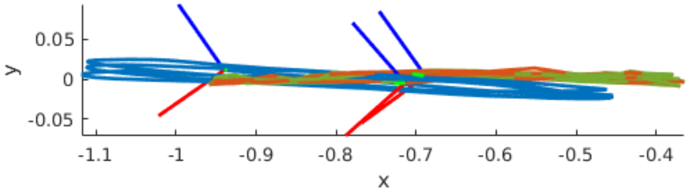
RESULTS: OXFORD MULTIMOTION

RESULTS: OXFORD MULTIMOTION



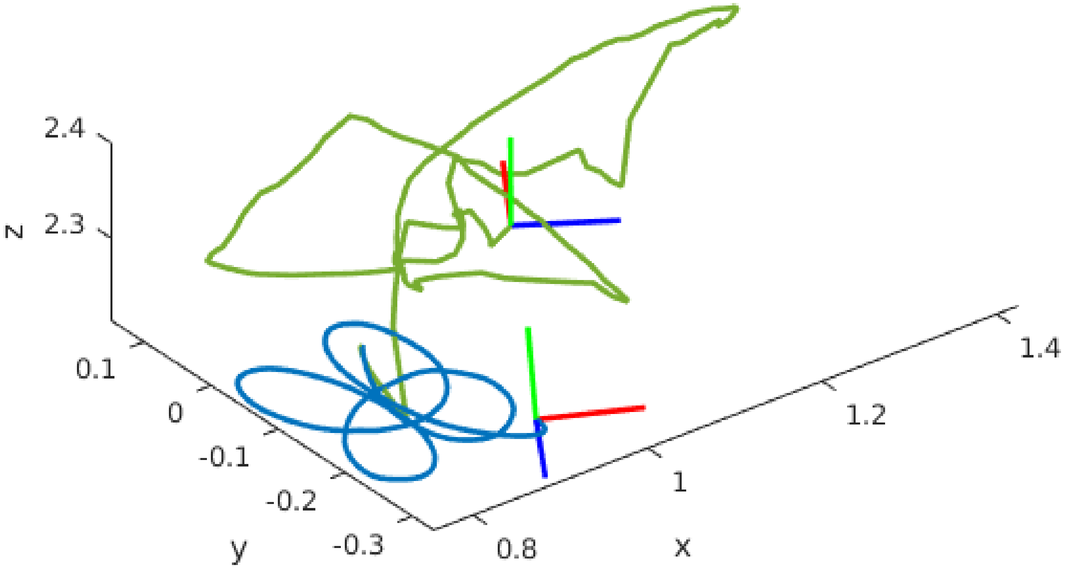


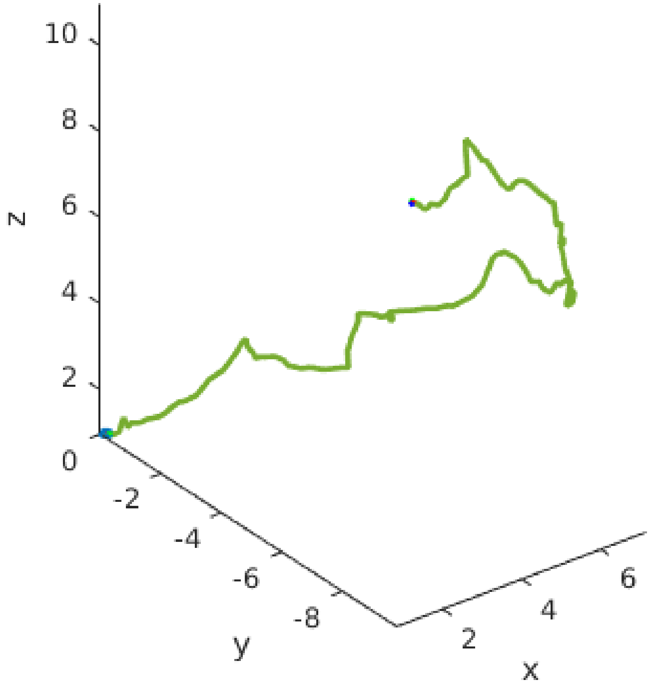




Good performance for translational motions.

RESULTS: OXFORD MULTIMOTION

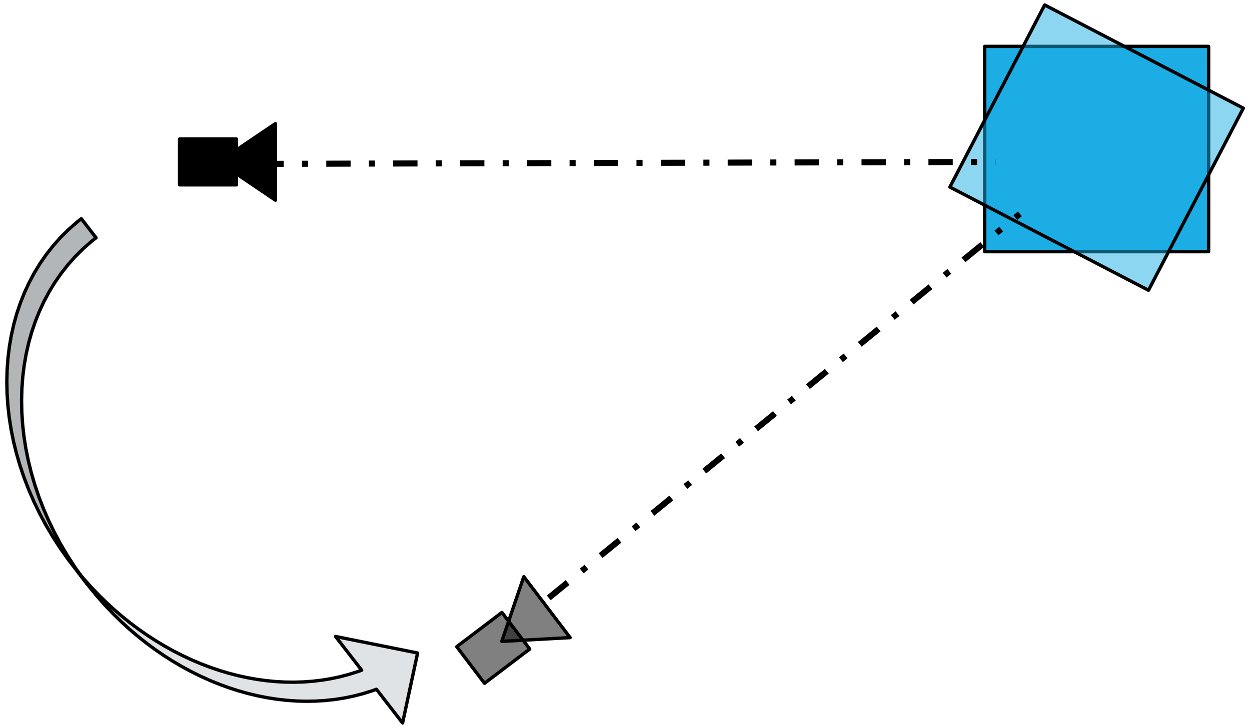




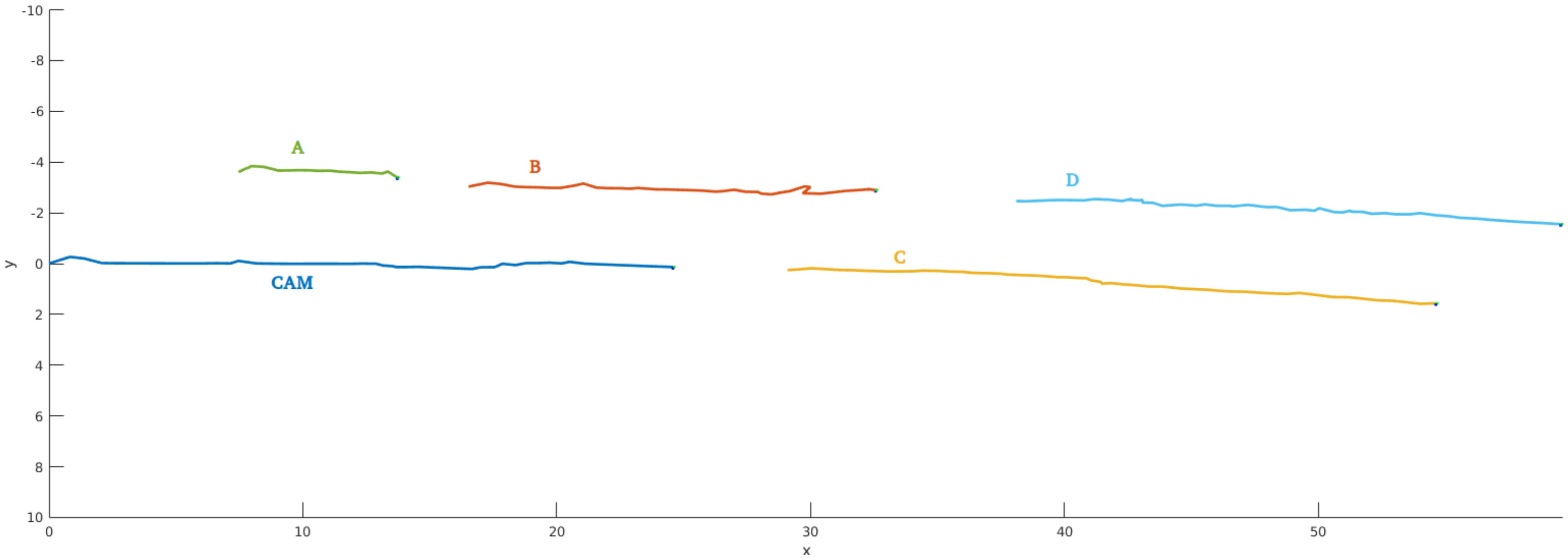
Poor performance for rotational motions.

RELATIVE ROTATIONS

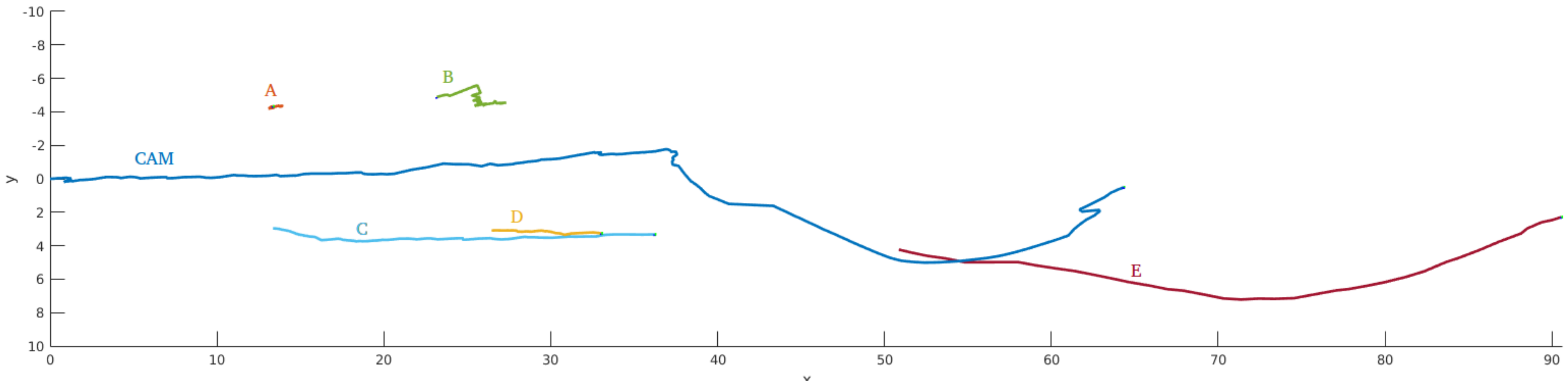
Hypothesis: Rotational motions induce large 'virtual' translation.



RESULTS: VKITTI



RESULTS: VKITTI



CONCLUSIONS & FUTURE WORK

- It is possible to track moving objects using dynamic photometric odometry, but there are limitations.

Future work should focus on:

- Determining the source of tracking failure in the rotating case.
- Refining estimates of the 3D structure of the dynamic objects.
- Improve computational performance by fully exploiting the parallelism of the problem.