Matchable Image Transformations for Long-term Metric Visual Localization

Lee Clement, Mona Gridseth, Justin Tomasi and Jonathan Kelly

UNIVERSITY OF TORONTO

STARS LABORATORY

Motivation

- Metric visual localization is challenging over the long term due to appearance change
- Experience-based mapping works, but requires a large database of experiences to operate
- Reduce the number of experiences by learning an image transformation that explicitly improves feature matching across appearance conditions

Approach

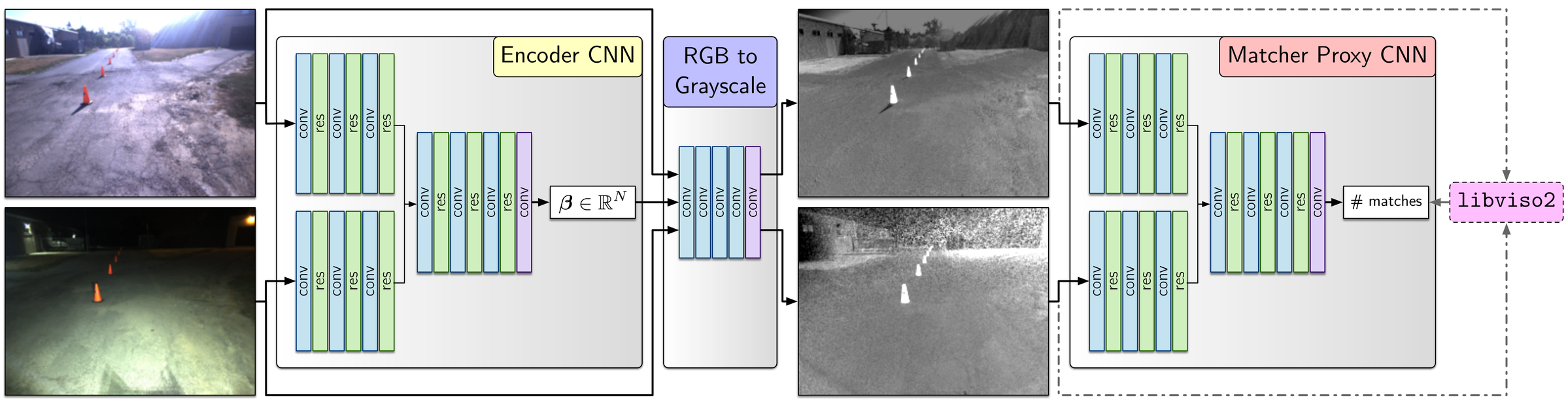
- Train a CNN to predict the performance of a non-differentiable feature detector/matcher
- Learn a maximally matchable RGB-to-grayscale mapping based on color-constancy theory

$$\mathbf{F} = \alpha \log \mathbf{R} + \beta \log \mathbf{G} + \gamma \log \mathbf{B}$$

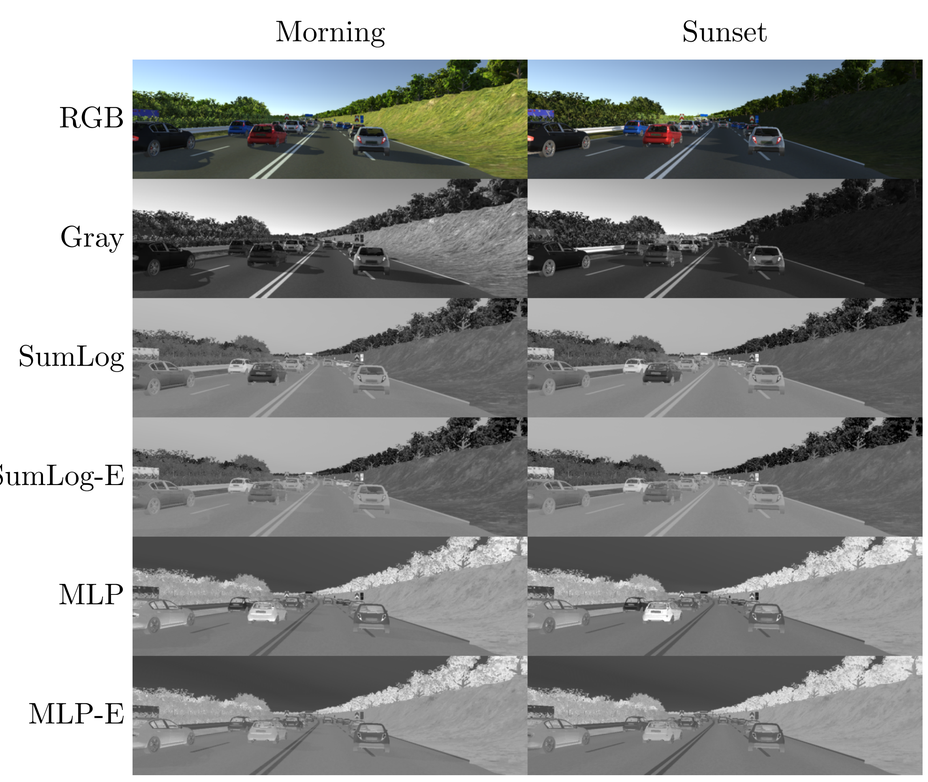
$$\boldsymbol{\beta} = \begin{bmatrix} \alpha & \beta & \gamma \end{bmatrix}^T = \mathcal{E}_{\boldsymbol{\phi}}(\mathbf{I}_1, \mathbf{I}_2)$$

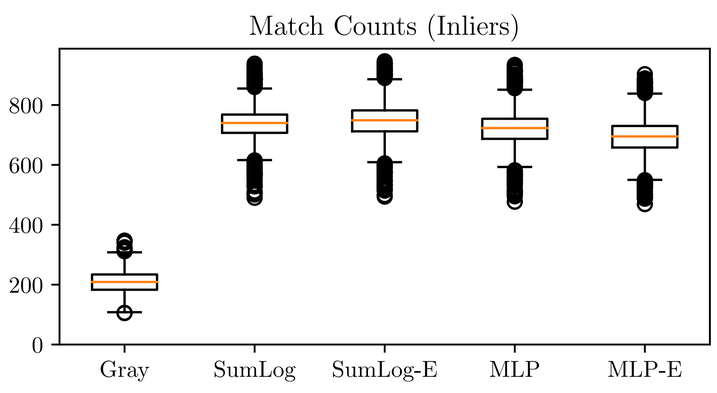
- Compare weighted sum-of-logs to MLP mapping

System Overview

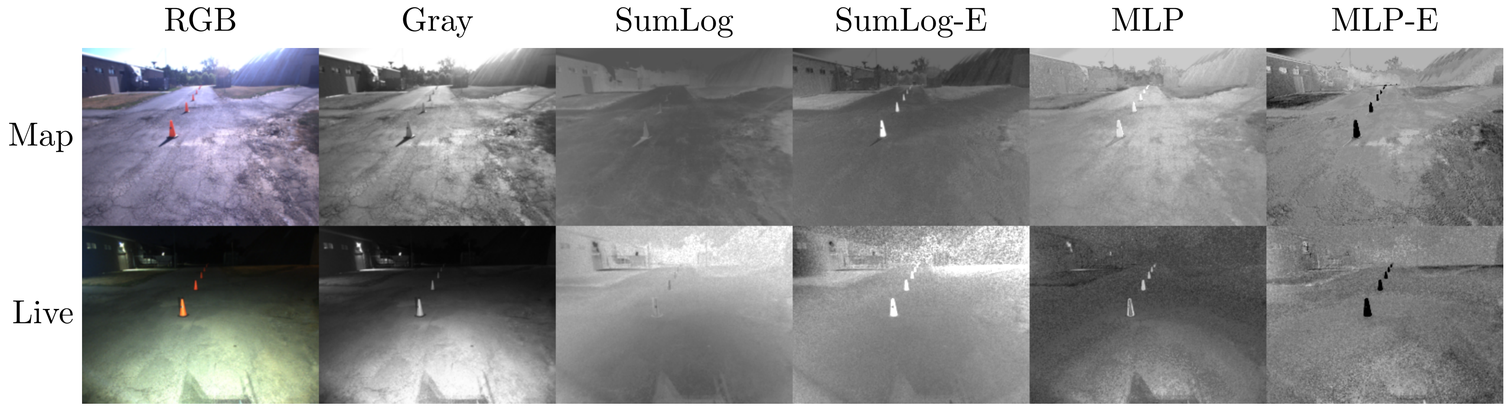


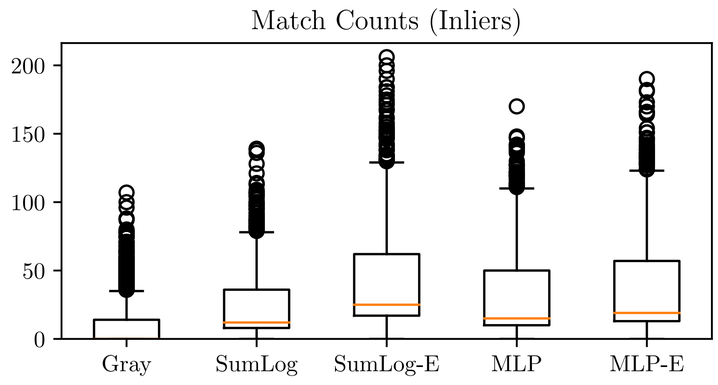
Results - Synthetic

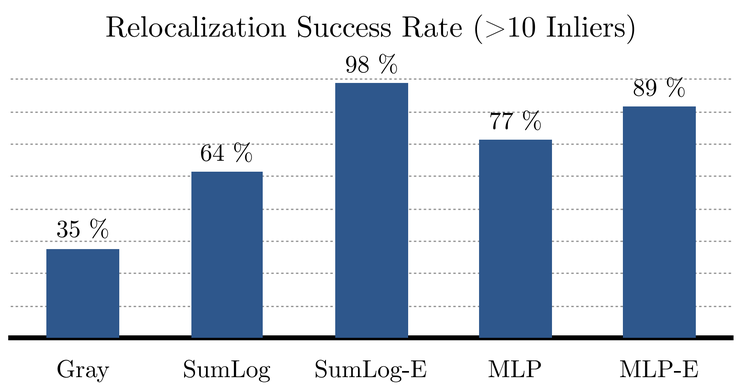




Results - Real







Conclusions

- Improved feature matching across day-night cycles
- Best performance achieved using physically motivated weighted sum-of-logs mapping with pairwise encoder CNN

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

starslab.ca

NSERC CRSNG NVIDIA VECTOR INSTITUTE