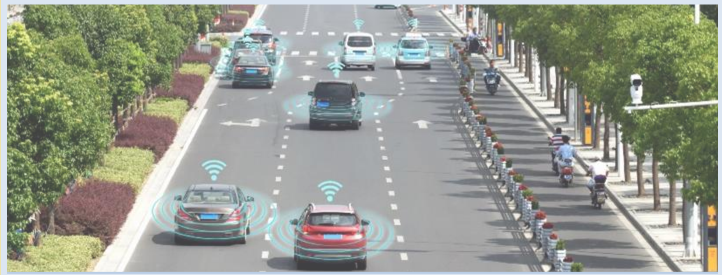
Deep learning-based covariance estimation for relative pose measurements

Alireza Ahrabian, Quan Nguyen, Nikos Toulios and Anthony Ohazulike

© Hitachi Europe Ltd. 2024. All rights reserved.

EVENTS Summary

EVENTS



Funded by the European Union

European Research Council

Objectives

- Development of robust and reliable perception of objects, and especially VRUs, under complex urban traffic and bad weather or low visibility conditions
- Improved perception performance while using cost-efficient sensor suites

https://www.events-project.eu/

EVENTS project has received funding under grant agreement No 101069614. It is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.

© Hitachi Europe Ltd. 2024. All rights reserved.

Introduction

Hitachi Inspire the Next

- Relative pose measurements obtained/derived from onboard vehicle sensors are often quite important for localisation and mapping tasks.

- The integration/fusion of such measurements can be used to obtained a vehicles pose in some “odometry” coordinate frame.

- Covariance estimates/measurements can be used to help improve accuracy via measurement weighting during estimation/optimisation.

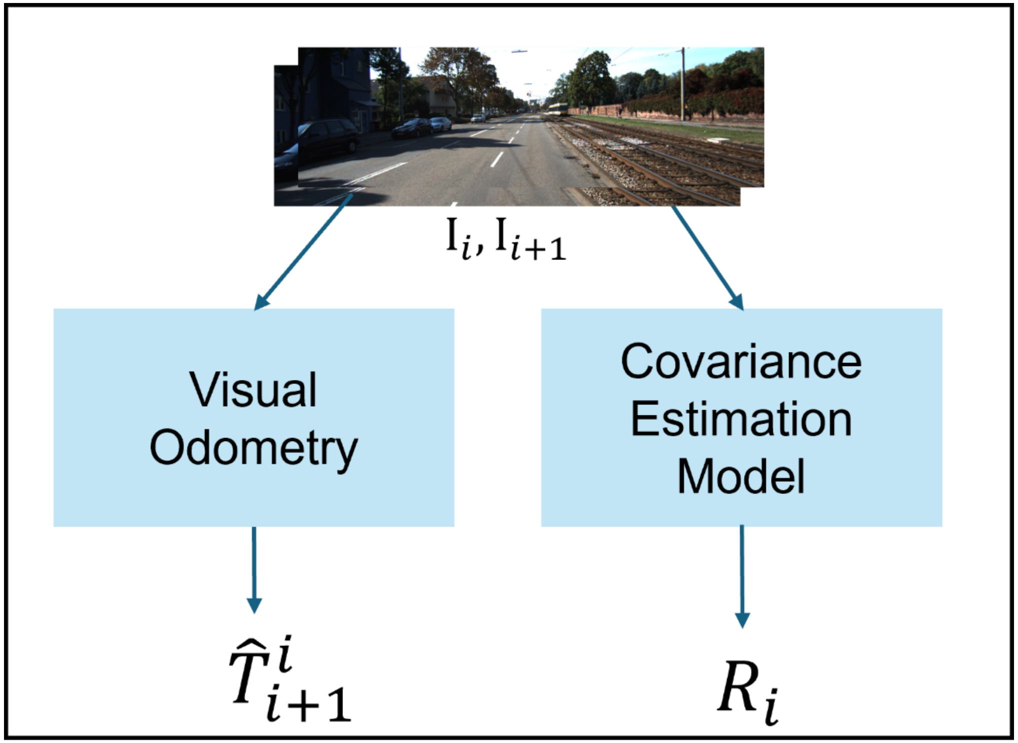
- Relative pose measurement errors can vary dramatically depending on scene/time (variation that is not Gaussian like).

- To this end, estimating covariance of relative pose measurements (with heteroscedastic errors) can potentially help improve accuracy of estimated vehicle pose.

© Hitachi Europe Ltd. 2024. All rights reserved.

Background – Existing Approaches

- In this work we build upon the methods proposed in the following:
  - K. Lui et al, “Deep Inference for Covariance Estimation: Learning Gaussian Noise Models for State Estimation”, 2018
  - A.D. Maio et al, “Simultaneously Learning Corrections and Error Models for Geometry-based Visual Odometry methods”, 2020.



© Hitachi Europe Ltd. 2024. All rights reserved.

Background – Covariance Estimation (Localisation)

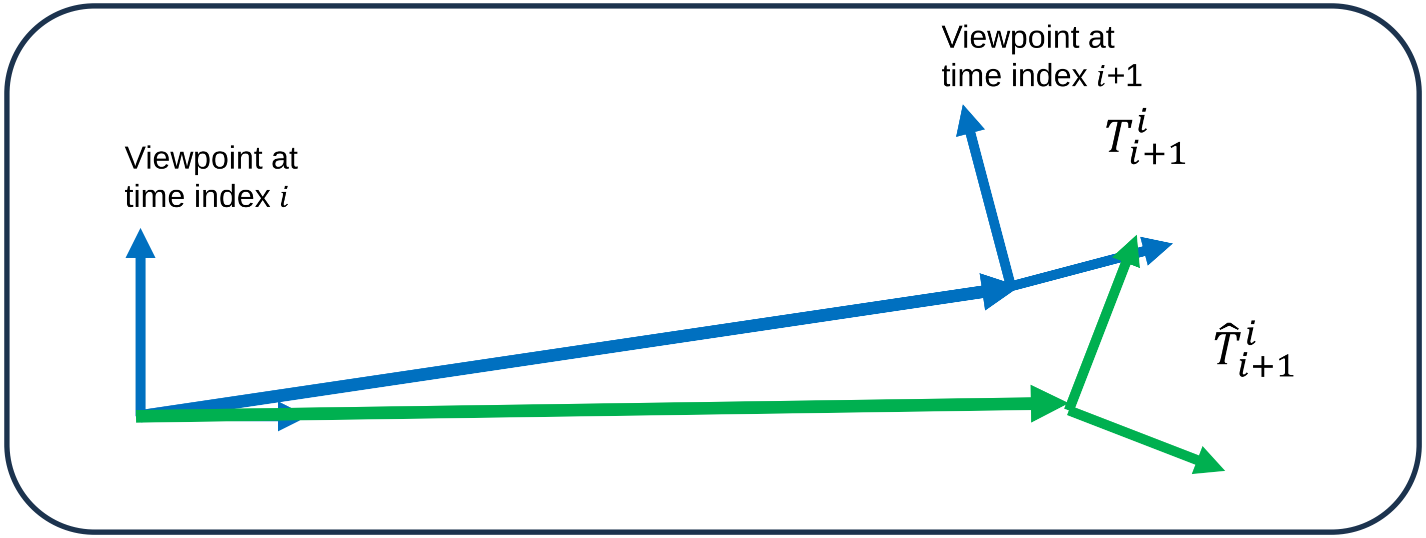
- The relative pose measurements ($\hat{T}_{i+1}^{i} \in R^{4\times4}$) are obtained from a visual odometry algorithm

- The error between ground truth relative pose (T_{i+1}^{i}) and estimated relative pose is given by:

$$\mathrm{e_i} = \log(\hat{T}_{i+1}^{i} T_{i}^{i+1})$$

- To this end, the covariance estimation models seek to determine the covariance R_i of the following error distribution:

$$\mathrm{e_i} \sim \mathcal{N}(0, R_i)$$



© Hitachi Europe Ltd. 2024. All rights reserved.

Limitations and Proposed Approach

- Our rational is that a diagonal estimation of covariance reflects the available information, and furthermore, and ultimately the diagonal covariance for a single measurement is the square of the error.

- To this end, our objective is to measure the squared error of the estimated relative pose and ground truth relative pose (error specified in previous slide).

- Prior work estimate covariance using a pair of mono or stereo images obtained from two different viewpoints.

- Prior approaches consume as input images obtained from two different viewpoints, therefore they are not able to estimate squared error for an arbitrary relative pose measurement.

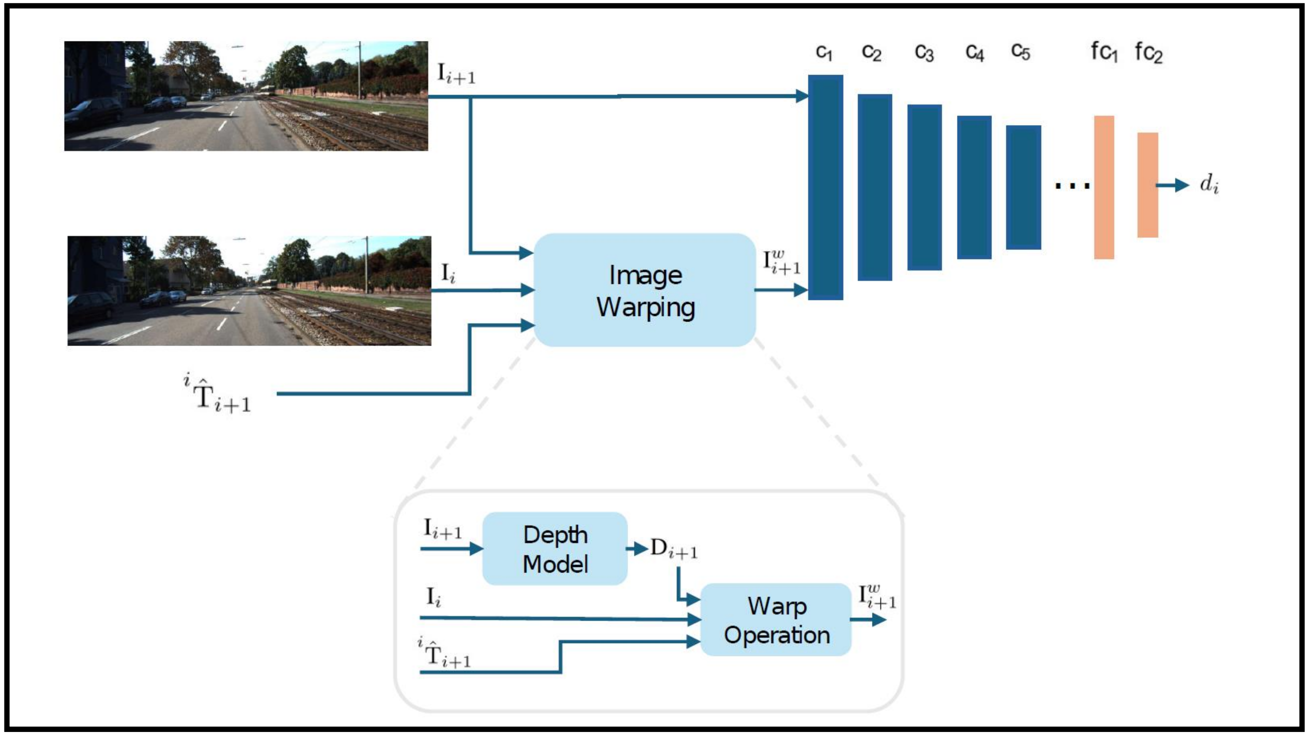
- Our proposed approach seeks to include the relative pose measurement as an additional input in order to capture error variation for a **general** relative pose measurement.

© Hitachi Europe Ltd. 2024. All rights reserved.

Proposed Approach

- The input consists of images acquired from two different viewpoints along with the measured relative pose, while the output is the estimated covariance.

- The model utilises the measured relative pose via an image warping operation, thereby enabling the covariance estimate to vary according to measured relative pose.



© Hitachi Europe Ltd. 2024. All rights reserved.

Training Details

- For the ***depth estimation network,*** we propose to utilise the pre-trained monocular depth network (monodepth2) and more specifically the 1024x320 version.

- As this is unscaled depth we use LiDAR measurements to compute a scale factor. Specifically using the ratio mean radial distance derived from LiDAR measurements and mean depth of unscaled depth estimates.

- The hyperparameters for covariance estimation network training were set as follows: batch size: 84, optimiser: Adam, and learning rate: 1e-04.

- Stoppage criterion: We stopped training according to that is, once we observed diverging train and evaluation losses.

© Hitachi Europe Ltd. 2024. All rights reserved.

Experimental Description

- We train and evaluate on the KITTI dataset. More specifically we use sequences 04-09 for training and sequence 10 for evaluation.

- Ground truth measurements were assumed to be the RTK-GPS measurement,

- In our work we considered relative pose measurements from the following “system”:
  - System 1: Sparse visual odometry
  - System 2: RTK-GPS

- Prior approaches would not be able to compute a covariance that considered the measured relative pose.

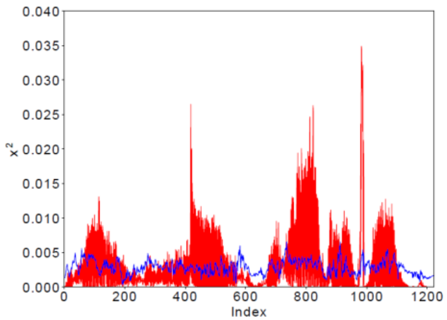
- We compute a variation of median absolute error between the ground truth squared error against the estimated squared error.

© Hitachi Europe Ltd. 2024. All rights reserved.

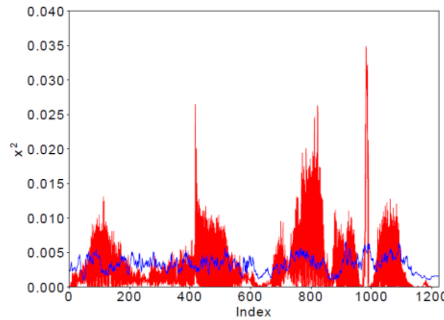
Results – Inference Examples

HITACHI
Inspire the Next

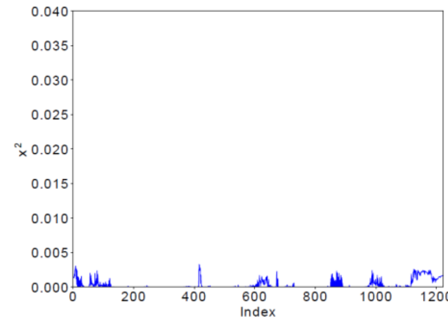
Translation - x



(a) Proposed method - System 1

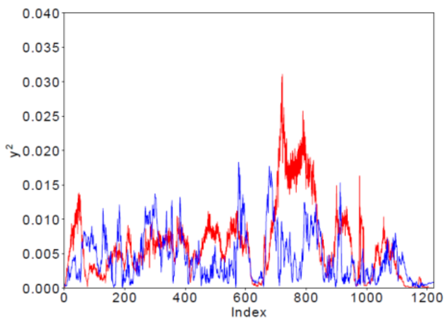


(b) [7] - System 1

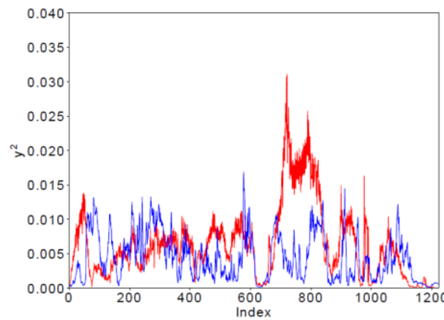


(c) Proposed method - System 2

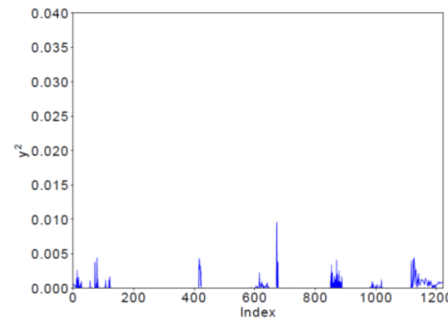
Translation - y



(d) Proposed method - System 1

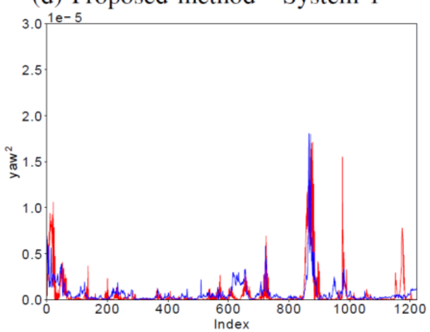


(e) [7] - System 1

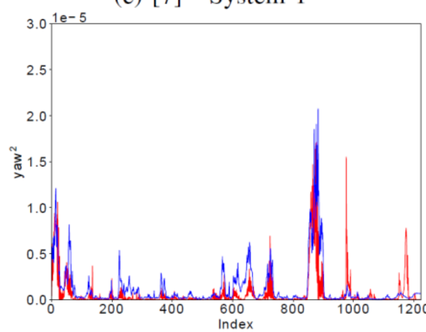


(f) Proposed method - System 2

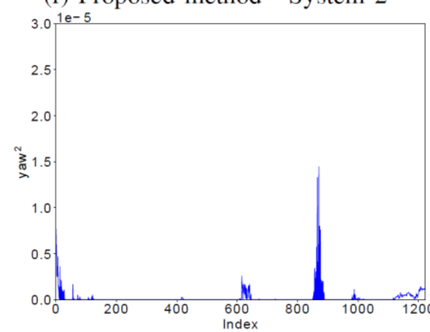
Rotation - yaw



(g) Proposed method - System 1



(h) [7] - System 1



(i) Proposed method - System 2

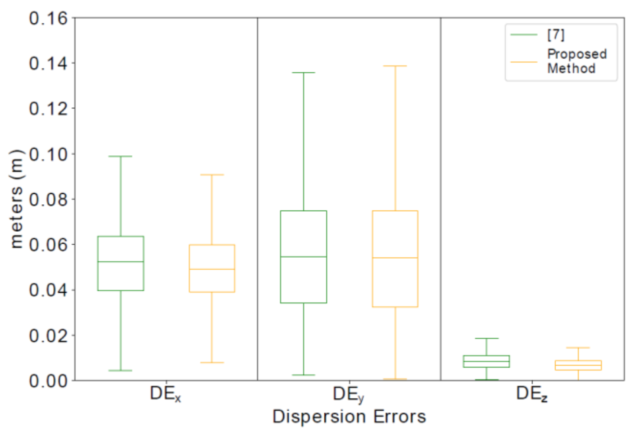
Ground Truth – Red line
Inference – Blue Line

[7] A.D. Maio et al, “Simultaneously Learning Corrections and Error Models for Geometry-based Visual Odometry methods”, 2020.

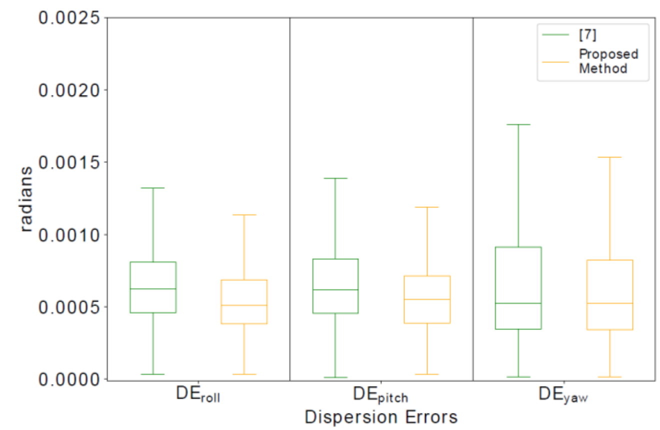
© Hitachi Europe Ltd. 2024. All rights reserved.

Results

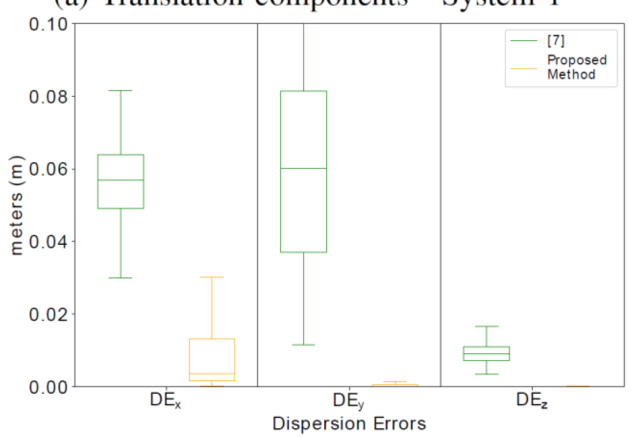
HITACHI
Inspire the Next



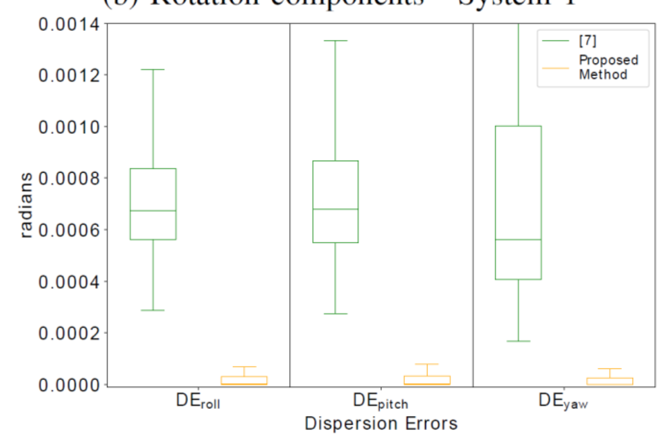
(a) Translation components - System 1



(b) Rotation components - System 1



(c) Translation components - System 2



(d) Rotation components - System 2

© Hitachi Europe Ltd. 2024. All rights reserved.

Conclusion

HITACHI
Inspire the Next

- In this work we have proposed a covariance estimation method for relative pose measurements obtained from an arbitrary system

- We achieve this by utilising the measured relative pose along with the images acquired between the two viewpoints to compute a covariance

- We demonstrate the ability of our model to estimate covariance for relative pose measurements obtained from two different systems

- Finally, in future work, we aim to investigate viewpoint synthesis via NeRFs as replacement to image warping, along with relative pose measurement generation to augment existing real-world datasets

© Hitachi Europe Ltd. 2024. All rights reserved.

END

© Hitachi Europe Ltd. 2024. All rights reserved.