Reply on RC2

Question 1: It should be noted somewhere (e.g. in the abstract) that the system cannot work at night.

The performance assessment of the infrared Raspberry Pi Camera Module 2 NoIR for capturing nighttime timelapse images was conducted by the authors. The Camera Module 2 NoIR shares all the features of the standard Camera Module, except for one distinctive attribute: the absence of an infrared filter, enabling the acquisition of images in low-light conditions using infrared illumination.

However, the experimental outcomes yielded inconclusive results. While definitive assertions regarding the system's nocturnal efficacy remain elusive, prudent exploration of alternative avenues is warranted before arriving at a final verdict. The following two potential courses of action merit consideration:

1. **Enhanced NoIR Cameras**: It is prudent to contemplate the utilization of superior quality NoIR cameras, particularly since Raspberry Pi's cost-effective offerings may not align optimally with specific requirements such as validating vision frameworks under low-light circumstances.
2. **Tailored Deep Learning Models for Nighttime Semantic Segmentation**: The development and application of deep learning models expressly tailored for nighttime semantic segmentation could conceivably bolster the camera's performance in challenging lighting conditions.

Notwithstanding these prospective enhancements, it is pivotal to underscore that the principal focus of this study was to introduce a holistic framework showcasing the viability of employing Computer Vision, Machine Learning techniques, and LiDAR data for the estimation of stream flow characteristics. Given this concentrated objective, the pursuit of further exploration into nocturnal image capture was regrettably deferred at this juncture.

Question 2: L142 The geographical coordinates of the site would be useful.

The selected case study location functions as a small-scale illustration, providing a testing ground for the proposed framework's assessment and result validation. It is important to note that the scope of this research does not encompass hydrological analyses pertaining to the catchment area. Consequently, the decision has been made to abstain from incorporating supplementary figures at this juncture, given the already substantial number of figures already presented in the study.

Question 3: 4.3 L266 Is the focal length fixed or variable?

Within the confines of this study, the Raspberry Pi Camera Module 2 was harnessed, notable for its unalterable focal length. The official documentation detailing the hardware specifications of this camera can be accessed via the following link: https://www.raspberrypi.com/documentation/accessories/camera.html.

Question 4: Why don't you estimate the DLT parameters directly? What is the advantage of estimating the intrinsic parameters beforehand?

The application of the Direct Linear Transformation (DLT) method mandates a minimum of 6 pairs of established 3D-2D correspondences. This prerequisite yields 12 equations, enabling the estimation of the complete set of 12 parameters inherent to the projection matrix. This matrix encapsulates both the camera's intrinsic and extrinsic parameters. Notably, intrinsic parameters remain invariant for a given camera model (e.g., Raspberry Pi Camera Module 2), enabling their

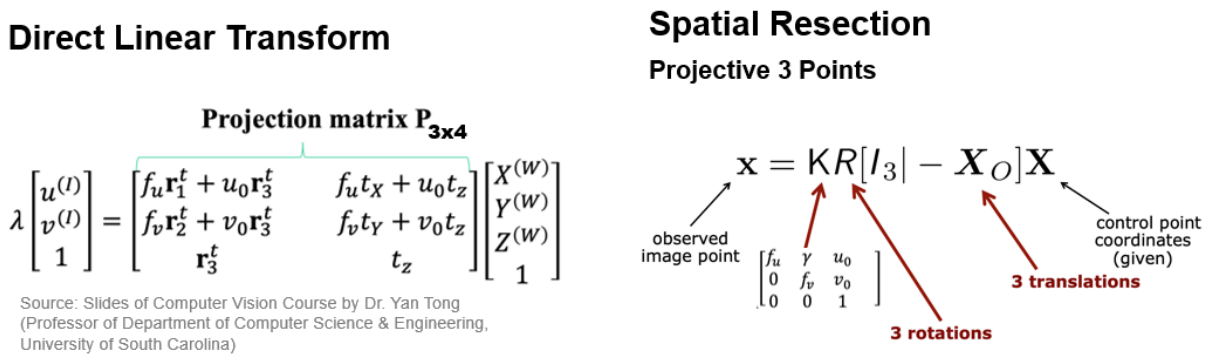

Figure 1. Direct Linear Transform method needs at least 6 pairs of known 3D-2D correspondence to calculate the projection matrix while Spatial Resection needs 3 pairs of 3D-2D points.

reuse across all images captured by that specific camera. Conversely, the extrinsic parameters undergo variation with shifts in camera position. Consequently, recalculating extrinsic parameters becomes imperative for each deployment scenario, ensuring the accurate reconstruction of the projection matrix.

To streamline the requirement for 3D-2D correspondence pairs, the Spatial Resection method was introduced. This algorithm adeptly capitalizes on existing intrinsic parameters, directing its focus toward the reevaluation of solely extrinsic parameters. By employing a mere minimum of 3 pairs of 3D-2D points, the Spatial Resection method successfully deduces the 6 unknowns essential for representing rotation and translation parameters (as illustrated in Figure 1).

Editorial:
Question 5: Eq. 1: camera (camera)

It has been fixed. Thank you!