Question:
What's the real challenge to reconstruct the outdoor garden where the trimming robot is driving?

Response:
Before understanding the challenges, we must take the real scenario and need into consideration. The 3D dense reconstruction uses the data from a driving robot in an outdoor garden with cheap image sensors (each costs 1 dollar, thus, low frame rate) in the Trimbot2020 project.

Challenge 1 from the Hardware

- The data is captured from an outdoor moving robot with fast speed (1 m/s and 90 deg/s).
- The data needs to be transferred to the remote server for 3D reconstruction by the mobile internet. The mobile internet speed is not fast and we would better downsample the video sequences (which will increase the rigid transformation between two adjacent frames) to save transferring time.
- The image sensor is required to be cheap (each one costs 1 dollar) in our project thus has a low performance (e.g. low frame rate).

All these requirements make the rigid transformation between two adjacent frames become bigger, which is a challenge to recover an accurate global pose trajectory.
To overcome this challenge, we developed a new device (panoramic stereo camera) and algorithm (Algorithm 1 in our paper).

Challenge 2 from the Application

- The aim is to reconstruct the surface of an outdoor garden. In an outdoor garden, the light and scenes tend to make the system sensitive.
- The 3D reconstruction result is used for users' remote visualization of the scene and robot task planning. Thus, the reconstructed point cloud must be dense and accurate to make its converted mesh vivid, which requires the estimated depth to be dense and accurate.

Feature-based SLAM methods (e.g. SFM, Orbslam) are sensitive to light change and can only produce sparse or semi-dense reconstruction results based on the matched features. By using the dense point clouds rather than the features (e.g. SIFT, ORB), we could eliminate this influence. Thus, the acquisition of dense point clouds or depth maps is the real challenge in an outdoor garden. ToF-based sensors do not work accurately outdoors because of infra-red light interference. Stereo vision cameras are not accurate and robust in real outdoor environments because of texture similarity, lighting changes, and shadows. Lidar sensors are expensive and their point clouds are sparse.
To overcome this challenge, we choose the image sensors with proper focal length, two complementary stereo vision algorithms, and the disparity fusion algorithm in the proposed SLAM system to improve the depth accuracy and robustness. Along with our proposed framework, a watertight 3D surface of the outdoor garden can be computed.

In conclusion, the real gardening application results in big transformations between adjacent frames and inaccurate depth information, which affect 3D dense reconstruction accuracy. As far as we know, there is no existing cheap device that can provide real-time 360 deg dense point clouds (lidar scanners are expensive while their data is sparse). And, there is no existing dense 3D reconstruction framework which could solve the challenges above properly. Thus, we have to first design the specific economical hardware (our panoramic stereo camera), along with the corresponding algorithms and datasets to solve the real challenges for the trimming robot in the real garden.