Live Demonstration: An Active System for Depth Reconstruction using Event-Based Sensors

Julien N.P. Martel*†, Jonathan Müller*, Jörg Conradt†, and Yulia Sandamirskaya*
*Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland
†Neuroscientific Systems Theory, Technical University of Munich, Munich, Germany

***Abstract*—We demonstrate a system that can reconstruct the three-dimensional structure of a scene using two Dynamic Vision Sensors (DVS) and an active component: a Mirror Galvanometer-driven Laser (MGDL). Our system uses two concurrent methods to estimate depth: 1) triangulation from the two DVS cameras calibrated as stereo rig, where stereo matching is greatly simplified by actively generating events in the scene with the laser beam, and 2) a technique derived from structured light, where we send the laser beam in a known direction and detect its projection in each of the two cameras. For the demonstration, we add two interactive components: the user can trigger (re)scanning of a specific location in the scene by pointing a hand-held laser pointer to it, and virtual objects can be rendered with correct scaling and distance on a screen using the reconstructed depth.**

I. Description of the System

Our system recovers the three-dimensional structure of a scene using event-based sensors and an active component: a Mirror Galvanometer-driven Laser (MGDL) [1]. The two event-based sensors are calibrated in a stereo-configuration and additionally, the MGDL is calibrated with each camera individually. Our setup is illustrated in Figure 1.

In this setup, we can estimate depth with two simultaneous methods: using active stereo-vision [1] and using principles of structured light. In the case of stereo-vision, we deflect the laser beam with the galvanometer-mirrors either randomly or systematically, e.g. in grid scanning. The deflected laser beam produces localized brightness changes in the scene, which can be detected as event blobs in the two DVS. Because of the high time resolution of the sensors, the blobs can be matched across the two sensors, which then allows triangulation.

Our second method uses the fact that the laser is also calibrated with each of the cameras individually. Thus, we know for any laser command we send (a command being the position of the mirrors deflecting the laser) in which direction the laser beam is pointing. Based on this direction and the projection of the laser point in the DVS, we can triangulate the 3D coordinates of the laser point.

In regions where the field of view of both cameras as well as the angular range of the laser overlap, both the stereo vision and the structured light methods can be used. In regions only visible from one camera (and accessible by the laser beam), only the method derived from structured light can be used. We can thus benefit from three depth measurements in the first case while a single measurement is available in the second.

II. Demonstration and User Experience

In our demonstration, we illustrate how these two active depth estimation methods perform and we add user interactions to exhibit features of our system: We provide two additional hand-held lasers (that we refer to as A and B) that can be manually pointed into the scene by the users. Each of these lasers blinks at a different frequency –also different from the frequency of the MGDL of the system itself– so that they can be uniquely identified and enable two "modes" of interaction.

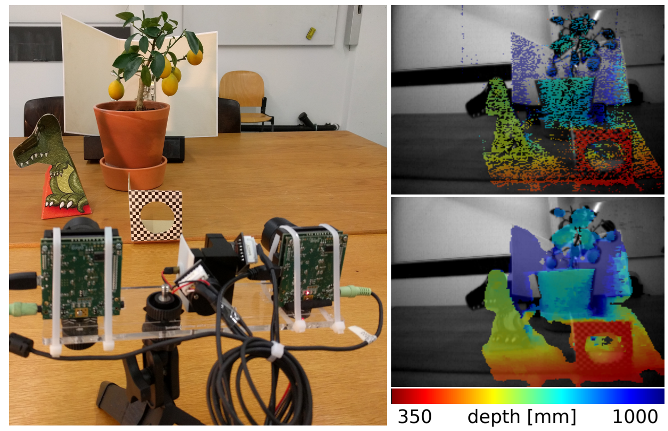

Fig. 1. **Left**: Stereo rig with two DVS and a MGDL in a scene. **Right**: Sparse and inpainted depth maps of the scene, reconstructed using our system.

In the "default-mode", i.e when no interaction with the user happens (that is, neither laser A nor B are detected), the system is continuously scanning the scene using its MGDL and reconstructing the 3D scene using the two concurrent depth estimation methods. The reconstructed scene is shown on the screen and is updated when the scene changes.

In a first interactive mode, the laser point created by the hand-held laser A is followed automatically by the MGDL by pointing its laser beam at the same 3D location. When laser A is detected in the same location for a certain time, a rescan in a small neighborhood around this location is triggered. In addition to demonstrating how depth scanning is performed, this mode shows how the calibration of the MGDLwith the two DVS allows to point the laser to an arbitrary coordinate in space.

In a second interactive mode driven by the hand-held laser B, the point it creates in the scene is detected and triangulated at high-speed. Since its three-dimensional coordinates can be fully estimated, we can pop-up virtual objects, correctly scaled, which can be seen on the screen as in virtual-reality.

III. Conclusion

Our demonstration aims to show how the event-based character of the DVS can be exploited in an active-vision setup for depth reconstruction. In interaction with the system, the participants can challenge and validate performance of the system.

References

[1] J. N. P. Martel, J. Müller, J. Conradt, and Y. Sandamirskaya, "An active approach to solving the stereo matching problem using event-based sensors," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, submission ID 2466.

978-1-5386-4881-0/18/$31.00 ©2018 IEEE