VSLAM and VLC Based Localization

David Vadeny, Mei Chen, Elaine Huang & Hany Elgala

University at Albany- State University of New York
College of Engineering & Applied Sciences, Computer Engineering
Albany, NY, USA

Abstract — **For this contest we propose to employ two complementary modes of localization to form a robust indoor positioning system. The first being computer vision based SLAM techniques for generating an accurate three-dimensional map and trajectory throughout the space the camera(s) have moved. The second being a VLC approach where positioning can be obtained through triangulation. We propose to utilize these techniques in tandem so the strengths of each technique will overcome the limitations of another.**

Keywords—Computer Vision, SLAM, VLC, Visible light, Direct SLAM

I. Introduction

The widespread use of smartphones and personal devices with at least one image sensor has enabled computer vision techniques to enter everyday lives. Applications range from face detection when taking a photo, car detection in traffic cameras, to even placing digital objects into what we see through virtual or augmented reality. With the use of cameras, we are able to give computers a way to perceive the space around them much the same how eyes give perception to humans. Leveraging the compute and storage capabilities of modern computers, perception can be extended beyond objects in a scene to a full three-dimensional model of the space.

The visible light communications (VLC) technology uses the visible light to transmit data while maintaining high-quality illumination. Light emitting diode (LED) based light bulbs or luminaires have the ability to be modulated at a rate higher than the imperceptible flickering range of a human eye. In optical camera communications (OCC) systems based on IEEE 802.15.7r1, image cameras are considered as a receiver [1]. Both the global shutter and the rolling shutter cameras are supported. Rolling and global shutter methods have their advantages and disadvantages [2]. As for the modulation technique, for example, the modulation of the transmitted signal includes On-OFF keying (OOK), where a bit 1 is transmitted as a higher intensity relative to the intensity used to transmit bit 0.

We propose a hybrid approach to indoor localization, taking advantage of two fields of techniques, namely dense simultaneous localization and mapping (SLAM) and visible light triangulation [3]. Dense SLAM is localization and mapping technique that has seen applications in virtual reality as well as 3D reconstruction. It is useful for mapping and localizing the room in which the sensor would be traveling through. Supplementary localization will be done through the triangulation of visible light with the use of custom LED based emitters deployed in the room. We believe these techniques in tandem constitute a robust indoor localization system.

II. Related Work

Many techniques have been proposed to tackle the VSLAM problem, and they typically rely on features, or robust sections of an image, that are unique enough to be easily located within a similar but different image of a scene. These features could be pixels or small clusters of pixels usually correspond to either corners of objects or blobs that easily distinguish themselves from most other portions of the image. An effective feature matching technique maps features from a new image back to the original image, thereby enables tracking of how these features have moved over time. This enables estimation of the translational and rotational motion of the camera, as well as that of the objects moving in the scene. Taking images in fast enough succession allows for accurate estimation of small movements of the camera, which then facilitates triangulation of the distance between objects in the scene and the camera, producing all the necessary components for a 3D model. Once a 3D model has been computed and stored, images can be taken with a new camera that is unaware of the scene and matched to a place inside the 3D model, allowing for localization.

Such techniques for solving VSLAM, however, come with an assumption and a weakness. The weakness is that VSLAM with a monocular sensor is subject to drift that builds up over time; the assumption is that the image features are sufficiently scattered throughout the scene and that they are relatively distinct from one another. This assumption is almost always satisfied in outdoor environments. Indoor environments, however, can be sparse in terms of salient features. This greatly reduces the reliability of feature-based methods and makes them unfit for indoor localization [4].

Advanced modulation techniques relative to OOK are also considered in OCC. In [5], binary frequency shift keying (BFSK) is considered to modulate the VLC signals by representing bits 1 and 0 using different frequencies. In [6], VLC for indoor positioning using cameras on mobile phones as receivers is considered. This research utilizes beacon to transmit data to a smartphone camera periodically taking pictures. OOK with custom beacon signals and a programmable microcontroller to control the LEDs are utilized. Bits of data that identifies the beacons are encoded with Manchester code.

III. METHOD

For the purposes of localization in an indoor space, we look to utilize this 3D model of the space to predict where the current image sensor is with respect to this model and provide an estimate of its location within the space. Computer vision problems to this end are known as Visual Simultaneous Localization and Mapping (SLAM or VSLAM). Solutions to this problem typically produce two useful pieces of information: the first being a model of the space that the image sensor has traversed though; the second, an estimation of the motion that the sensor has gone through once it began recording data.

To build a robust indoor localization system, we consider the use of the featureless VSLAM technique that is known as Direct SLAM [7] that makes use of the raw intensities of pixels within the image instead of features. This method promises to provide more geometric information about the scene as it operates on every pixel and not just a small set of features. It potentially relaxes the constraint on needing the environment to be distributed with unique points of contrast and thereby enables the system to function even within largely empty indoor spaces.

The Direct SLAM technique alone, however, is susceptible to small amounts of drift and jumps over regions for which it is not able to resolve. Pointing the image sensor at a large blank wall for example will reduce localization accuracy as nothing in the image frame is changing. Such situations call for complementary localization techniques that can provide estimation in this scenario. We choose to leverage the recent development in visible light positioning.

One of the main techniques of visible light positioning is triangulation, where exposure of a light sensor to multiple emitters simultaneously can provide the position of the receiver based on the known positions of at least three emitters using received signal strength measure (RSS). We plan to deploy multiple LED fixtures onto the ceiling, so as to be able to localize our experimental device even when the device is close to a wall and its vision no longer facilitates localization. Using this technique simultaneously with computer vision can correct drift that builds up from using Direct SLAM alone.

A noted limitation of visible light triangulation is that its accuracy can be greatly diminished when the receiver approaches the corners of a room due to the increase in noise [7]. Incidentally this weakness of triangulation can be easily addressed by the computer vision component because corners of a room typically present changing scenery. Therefore, visible light positioning can be effective when the ability of computer vision diminishes, and computer vision can produce a solid model in locations where visible light techniques tend to fail, working in tandem they would provide a robust solution to indoor localization.

IV. IMPLEMENTATION

For deployment we plan to use 5 custom made LED emitters for the purposes of visible light triangulation as well as a dedicated image sensor always pointed in the direction of these light sources. A microcontroller will be added to provide additional circuitry to allow data transmission, reduce flickering, and other functions.

For the computer vision side, we plan to use either modern cell phone cameras or hobby robotics cams. All images will be streamed to a nearby laptop for processing in a MATLAB environment for 3D modeling, localization, and triangulation.

References

[1] Rajagopal, Sridhar, Richard D. Roberts, and Sang-Kyu Lim. "IEEE 802.15. 7 visible light communication: modulation schemes and dimming support." Communications Magazine, IEEE 50.3 (2012): 72-82.

[2] Pathak, P. H., Feng, X., Hu, P., & Mohapatra, P. (2015). Visible Light C ommunication, Networking, and Sensing: A Survey, Potential and Chall enges. Communications Surveys & Tutorials, IEEE, 17(4), 2047-2077J.

[3] Engel, J. Sturm, D. Cremers “Semi-Dense Visual Odometry for a Monocular Camera”, In IEEE International Conference on Computer Vision (ICCV), 2013.

[4] J. Engel, T. Schöps, D. Cremers, “LSD-SLAM: Large-Scale Direct Monocular SLAM”, In European Conference on Computer Vision (ECCV), 2014.

[5] Rajagopal, N., Lazik, P., & Rowe, A. (2014, September). Hybrid visible light communication for cameras and lowpower embedded devices. InPr oceedings of the 1st ACM MobiCom workshop on Visible light commu nication systems (pp. 33-38). ACM.

[6] Kuo, Y. S., Pannuto, P., Hsiao, K. J., & Dutta, P. (2014, September). Lu xapose: Indoor positioning with mobile phones and visible light. InProce edings of the 20th annual international conference on Mobile computing and networking (pp. 447-458). ACM.

[7] Chunyue, Wang, Lang,Wang, et. al “The Research of Indoor positioning based on Vislible Light Communication” in Communications, China, vol 12, 2015, pp. 85 – 92.