https://doi.org/10.2352/ISSN.2470-1173.2020.13.ERVR-338

Heads-Up Lidar Imaging with Sensor Fusion

Yang Cai, Sean Hackett, Ben Graham, Florian Alber, and Mel Siegel
Carnegie Mellon University, Pittsburgh, PA Email: ycai@cmu.edu

Abstract

In this paper, we present a novel Lidar imaging system for heads-up display. The imaging system consists of the one-dimensional laser distance sensor and IMU sensors, including an accelerometer and gyroscope. By fusing the sensory data when the user moves their head, it creates a three-dimensional point cloud for mapping the space around. Compared to prevailing 2D and 3D Lidar imaging systems, the proposed system has no moving parts; it's simple, light-weight, and affordable. Our tests show that the horizontal and vertical profile accuracy of the points versus the floor plan is 3 cm on average. For the bump detection the minimal detectable step height is 2.5 cm. The system can be applied to first responses such as firefighting, and to detect bumps on pavement for low-vision pedestrians.

Introduction

Lidar (Light Detection and Ranging) is a remote sensing method that measures distance to a target by illuminating the target with pulsed laser light and measuring the reflected light with a receiver. Lidar has been widely used in satellite remote sensing systems for surveying terrane of the Earth. Lidar has several unique advantages: it can travel through vegetation and thin smoke for a long distance; its pulsed signal saves energy for the device; its coded pulse patterns enable the detection that is resilient to ambient noise. Recently, motorized sweeping multi-line Lidar has been widely used in autonomous driving vehicles to survey the 3D environment surrounding the vehicle and detect obstacles in real-time. It has become an essential sensor for autonomous driving vehicles.

So can a Lidar sensor be used for heads-up display applications such as on a firefighter's helmet? It is desirable that the first responders are able to scan the surrounding environment while walking inside a burning building. However, motorized Lidar sensors are expensive, heavy, and produce annoying vibrations. In short, it is not comfortable to have a spinning vibrating device on the top of your head.

Here we propose a heads-up Lidar system based on sensory fusion of a stationary one-dimensional Lidar with motion sensors, including accelerometers and gyroscopes. The one-dimensional Lidar is used for measuring the distance between the front end of the helmet and the target. The gyroscope is used for measuring the angle between the Lidar and the target. We use a gyroscope instead of magnetic field sensors as the gyroscope only needs to be calibrated once during the development stage, rather than calibrating the magnetic field sensors every time the helmet is used. There are a few automatic self-calibration designs for magnetic field sensors such as the thermal imaging drone Parrot Anafi [1], in which a motorized calibration wheel is included, but in our case, manual or motorized rotation both appear to be inadequately effective.

The heads-up Lidar system is virtually a 3D imaging sensor. It can generate a low-density point cloud by moving the head around. It's accuracy depends on the accuracy of the Lidar (distance) and gyroscope (angle). In this study, we focus on two applications: horizontal and vertical 2D imaging by sweeping the head from side to side/ up and down, and bump detection by pointing the Lidar to the ground while walking forward. This has several advantages: it has no moving parts; it is light-weight and affordable; it has the heads-up display to show the 2D or 3D map in real-time; and it can be embedded into wearable devices such as a helmet and can be reconfigured for multiple applications for example, user activity recognition [2]. However, it has its disadvantages as well: it is less accurate than an integrated 3D Lidar scanner; it has sparse 3D point sampling; and its imaging field is limited by the head movement.

Compared to other ranging mode systems, we found the proposed heads-up Lidar system is more promising for first response applications. Radar imaging systems can see through smoke and several types of walls but their imaging resolutions are low [3]. In many cases, the devices are too bulky to be mounted on the helmet. Ultrasound ranging systems have limited imaging distance through open air and their resolutions are low as well [4]. For the last two decades, Structure-from-Motion (SfM) has been evolved to a popular technology for 3D imaging with an affordable single camera, a pair of stereo cameras, or multiple cameras [5]. The RGB camera based SfM methods commonly need structural features such as Difference of Gaussian (DoG) SIFT features [6-7] or FAST corner features [8] to match the structural features between frames in the video and calculate the homographic transformation matrix accordingly for Simultaneous Localization and Mapping (SLAM) [9]. The matching algorithm needs a minimal number of features in consecutive frames of the video. In many cases, unfortunately, there are not enough matching features between frames, due to "featureless" smooth walls, blurry images, or fast movement of the camera.

On the other hand, studies show that the accuracy of SfM can be improved with multi-modal sensory fusion such as motion sensors (IMU) on a drone [10]. Multi-modal sensory fusion requires less computational power so that it can perform in real-time. With rapidly declining cost of sensors, thanks to mass production of mobile phones and robotic parts, sensory fusion enables a broader spectrum of reconfigurable applications and provides redundancy to perceptual intelligence, for example, the altitude of a user's helmet can be derived from a pressure sensor and a temperature sensor; and it can also be derived from IMU sensors, as well as GPS signals when they are available. Multimodal sensory fusion does not require rigid feature detection and feature matching. Instead, it is based on the first principles of physics and visualization of digital "pheromones" [11] along with the moving trajectories. In nature, pheromones are used for insects to communicate among each other and to map their environments. It is less computationally complex but requires more sensory fusion from live data. We anticipate this biomimicry approach would project a new direction for mapping, positioning, and navigating technologies [12].

Overlaying graphical and textual contents on heads-up displays (HUD) has been common in augmented reality (AR) systems, such as Google Glasses [13], Microsoft HoloLens 2 [14], and Magic Leap [15]. Most prevailing AR systems project virtual world models onto the real-world background, creating mixed reality experience such as the non-HUD game Pokémon on mobile phones [16]. In our study, we want to project live and on-demand information on the HUD for situational awareness, navigational instructions, and decision-making assistance. We call it "hyper-reality" or "super reality" technology, which display real-time information that is more than our eyes can see, such as thermal images, WiFi electromagnetic field distribution, objects behind the wall, victim in the smoke, and hazardous gases.

System Architecture and Global Coordinates

The heads-up Lidar sensory system contains a 1D Lidar for distance measurement up to 40 m with sampling rate up to 500Hz, a 10 DOF IMU sensor, including a 3 axis gyroscope, a 3 axis accelerometer, a 3 axis magnetic compass, and an altimeter. The helmet system includes a quad core GPU processor and a projection heads-up display. The sensors are placed on the front brim of the fireman's helmet. The 1D Lidar and 10-DOF IMU components are shown in Figure 1, and the helmet prototype in Figure 2.

Figure 1. WaveShare 10-DOF IMU sensor (left) and time-of-flight 1D Lidar distance sensor (right)

The Lidar direction and relative axis of the accelerometer and gyroscopes are shown in Figure 3. To fuse multiple sensory data elements, we need to convert the spherical polar coordinate system of the sensor outputs into a cartesian coordinate system. The standard relations between cartesian, cylindrical coincide with the positive polar axis of the spherical system coincide with the positive y-axis is illustrated in Figure 4. Then the conversion between spherical coordinates (θ, φ, p) and cartesian coordinates (*x, y, z*):

$$r = p \cdot sin\varphi \tag{1}$$

$$x = r \cdot cos\theta; \tag{2}$$

$$y = r \cdot sin\theta \tag{3}$$

$$z = p \cdot cos\varphi \tag{4}$$

Figure 2. The Hyper-Reality Helmet prototype that the sensors are embedded into standard firefighter's helmet

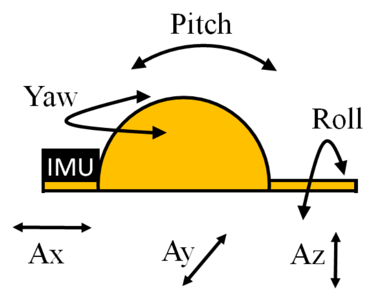

Figure 3. The position of 3-axis Accelerometer, 3 axis Gyroscope and Lidar in relation to the helmet

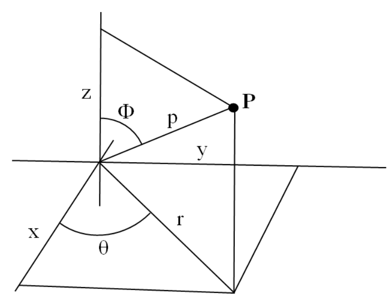

Figure 4. Relations between cartesian, and spherical coordinate system

Horizontal Sweeping 2D Imaging

By sweeping the head from side to side, a two-dimensional scan of a space can be obtained. This can be transformed into 3D cartesian space by converting the spherical coordinates from the Lidar distance, l, pitch, φ, and yaw, θ, to the cartesian coordinates. The position of the Lidar in relation to the true origin must also be taken into consideration as the distance from the centre of the head's rotation will be greater than the distance from the edge of the helmet where the Lidar is placed, as shown in Figure 5. The distance from the Lidar to the center of the head's rotation is denoted as , c, and shown in Figure 6. This activity is performed while stationary, and it is assumed no movement other than rotation of helmet takes place.

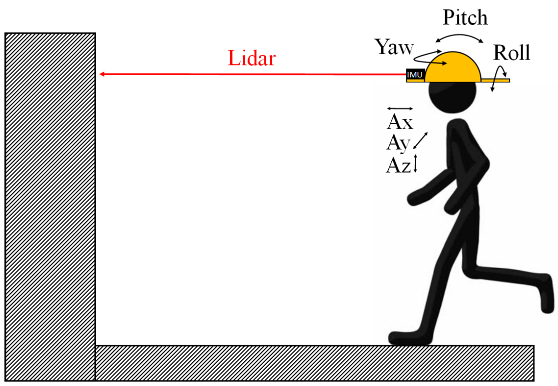

Figure 5. Horizontal sweeping 2D imaging: the profile view

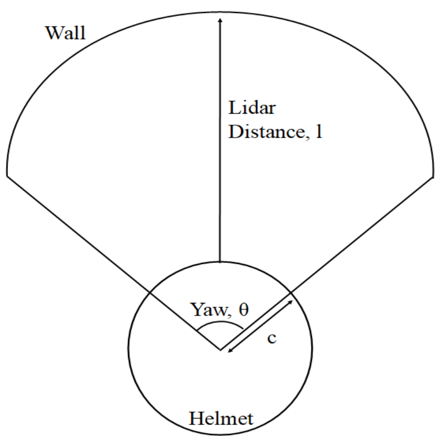

Figure 6. The bird's eye view of the sensory ranges

Pseudo code for Horizontal Sweeping 2D imaging:

```
-Import raw Helmet Sensor Data
-Adjust Lidar distance to the centre of helmet
-Convert spherical coordinates to cartesian
     using equation (1)-(4)
-Plot Lidar points in 2D and 3D space
-Add walls to plot obtained from building
     floor plans
-Calculate max and average error between floor
plans and Lidar points
```

Figure 7 shows screenshots of the scanned 2D map from the heads-up display. The point cloud on the top is the map and the grid at the bottom is the reference plane. The map is generated by sweeping the head from side to side.

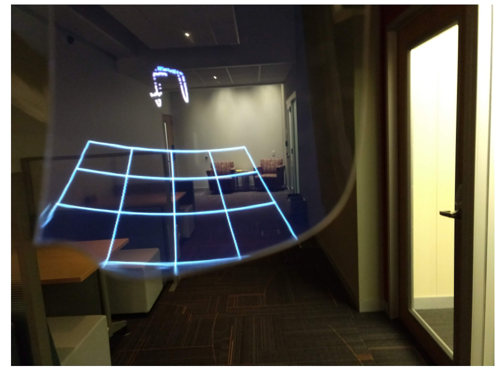

Figure 7. Screenshots of the scanned 2D map from the heads-up display. The point cloud on the top is the map and the grid at the bottom is the reference plane. The map is generated by sweep the head from side to side

Figure 8 - 11 shows 3 results obtained from the horizontal head sweeping in an office environment with different floor plans. It can be seen that some deviation results from the actual floor plan due to; small surface changes such as doors/door frames etc, the rotation of the helmet not having been done perfectly around the origin, and the drift of the gyroscope over longer periods of time. The average resultant error was typically ± 3 cm.

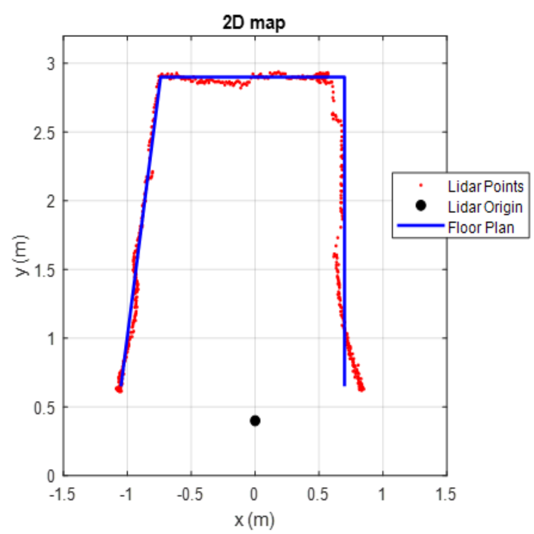

Figure 8. The result of the Lidar imaging for a U-shape space in comparison with the floor plan

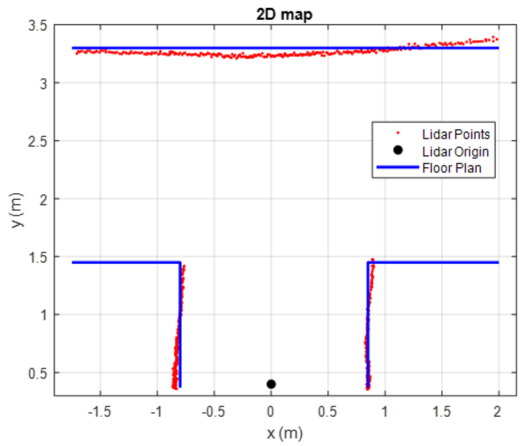

Figure 9: The result of the Lidar imaging for a T-shape space in comparison with the floor plan

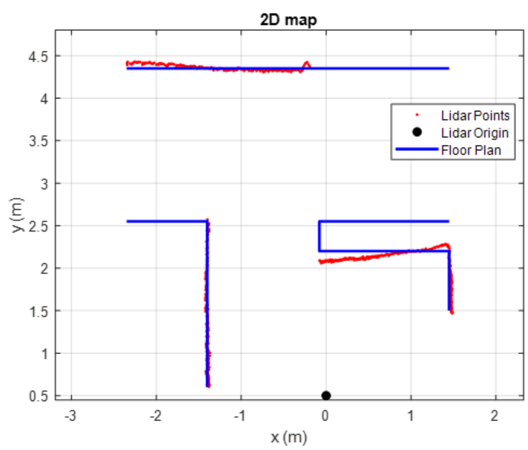

Figure 10: The result of the Lidar imaging for a complex space in comparison with the floor plan

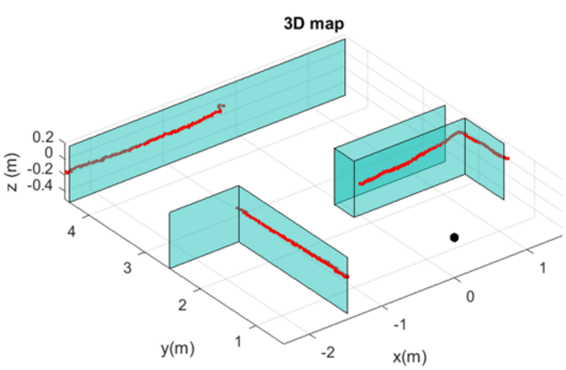

Figure 11. The 3D representation of the Lidar imaging for a complex space

Vertical Profiling - The "Virtual Cane"

For bump/step detection the Lidar can be positioned at a downward angle facing the ground. The height of the helmet above the ground can then be calculated, and thus any step changes detected. As the height of the user, h1, pitch, φ, Lidar distance, l, and distance between Lidar and head, c, are known the resultant step height can be calculated, as shown in Figure 13. On flat ground this should remain zero. The change in height of the helmet, h2, due to movement while walking also needs to be considered, and can be determined through the use of the 3 axis accelerometer, like displayed in Figure 12.

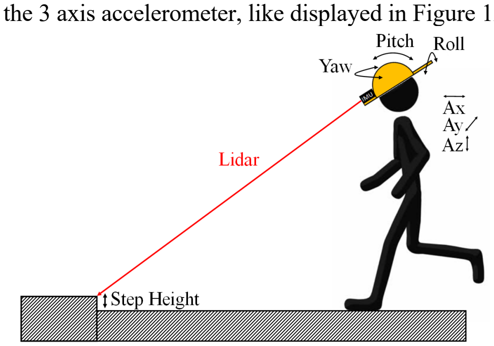

Figure 12: Vertical profile measurement

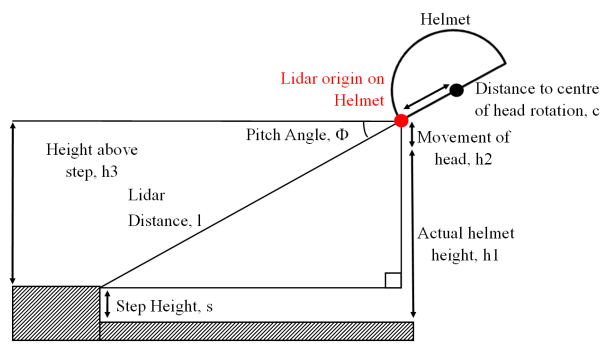

Figure 13. The relationship between helmet orientation and step height Lidar

This is achieved by applying a Kalman filter using the accelerometer and gyroscope data to improve the accuracy, removing the bias due to gravity and performing double integration to find the displacement of helmet.

The step size is calculated through the following process, beginning with the calculation of the helmet height above the ground, h3:

$$h3 = (l + c) \cdot sin\varphi \tag{5}$$

Then the displacement in the *x* and *z* axis of the accelerometer:

$$xS = \int_0^t \int_0^t xA \, .dt \, .dt \tag{6}$$

$$zS = \int_0^t \int_0^t zA \, .dt \, .dt \tag{7}$$

Then the resultant displacement in the *x* axis corrected for the pitch of the helmet:

$$h2 = xS \cdot cos\varphi + zS \cdot sin\varphi \quad (8)$$

So finally, the detected step size is:

$$s = (h1 + h2) - h3 \quad (9)$$

Like the horizontal sweeping imaging the same action can be formed in the vertical plane. By sweeping the head up and down while stationary, a scan of the space can be obtained as shown in Figure 14.

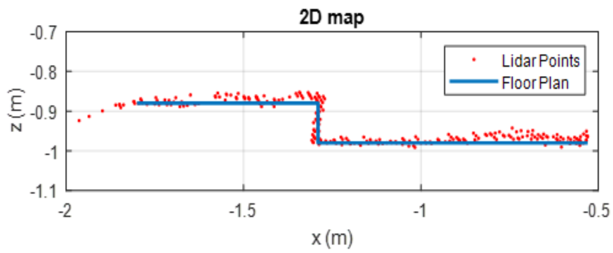

Figure 14. The result of the Lidar imaging for a step elevation in comparison with the floor plan

Pseudo code for bump detection;

```
-Import raw sensor data
-Data Processing
-Process accelerometer data to find
     displacement
-Use Kalman filter to improve pose based on
     accelerometer and gyroscope
-Remove bias due to gravity
-Perform double integration to get velocity
     and displacement
-Result for head movement, h2, obtained
-Step Detection
-Calculate step height, s, using the equations
     (5)-(9)
-Creating sliding window of step height
     average
-Calculate gradient between sliding window
     samples
-If gradient exceeds minimal detectable value,
     BUMP DETECTED
```

Bump Detection Results

Step change of the floor height are calculated as described in the Section about the Vertical Profiling - The "Virtual Cane". A plot of a bump on the ground detected while walking with the helmet is shown in Figure 15 and 16. The plot still contains some rapid changes in step height due to noise and the sensor range accuracy. This makes detecting the step more difficult. To overcome this a sliding window average of the step height is created for easier and more reliable identification of the step change. Each sliding window contains approx 0.5 seconds of data. The gradient between sliding window samples can then be calculated and steps can be easily identified if it detects a gradient above the minimum threshold. The minimum detectable level achieved is approximately 2.5 cm. This can then be labeled on the original step data, with red representing an up-step, and green a down-step.

While the accuracy of the vertical and horizontal scanning was most concerned with the drift of the gyroscope over time as the scan was taken, the bump detection is most concerned with the accuracy over a short time frame i.e. due to the accelerometer, as it is the change in step height from the previous sample that is being detected. This necessitated the use of the Kalman filter to get an accurate acceleration and thus displacement.

Compared to a visual camera bump detection system which may easily locate a step up based on e.g. shadows, it may have much more difficulty for a step down which is not as easy to visually identify. This Lidar solution does not suffer from this problem and is just as effective for detecting up and down steps. It also has much simpler computation than a visual camera detection system which would require more intensive image processing.

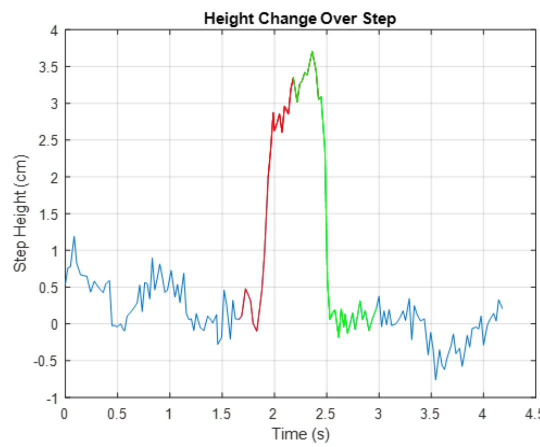

Figure 15. The vertical height changes over a step

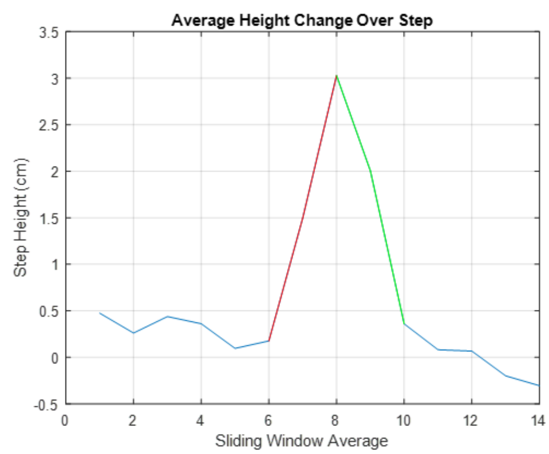

Figure 16. After sliding window average applied

Conclusions

In this study we have shown that sensor fusion of 1D Lidar, 3 axis accelerometer and 3 axis gyroscope can be used to accurately map out spaces in 2D and 3D. The accuracy of the helmet for horizontal and vertical profiling was able to achieve an accuracy of ±3 cm for the entire scan. For the bump detection solution, which calculated the step change in height along the ground, the minimal detectable step size between samples was 2.5 cm, which is sufficient for detecting the majority of steps which could cause a tripping hazard for low-vision pedestrians. Both up and down steps are detectable.

Discussion

There is still room for improvement with this methodology. For example, the movement of the Lidar on the helmet can result in errors in the step height calculation. This is reduced by use of the accelerometer to detect the up and down movement of the head but there will always be a certain accumulation of error when performing the double integration to obtain the displacement, especially when trying to detect such small displacements. A solution to this could be to mount the Lidar on a gimbal so it remains almost completely stationary. This though would have the disadvantage of added weight on the helmet which could result in discomfort.

The directionality of the 1D Lidar limits the helmet to detect steps in a direct line ahead. Any obstacles to the side could be missed. A 2D Lidar solution to counteract this could be employed, but will come with the associated disadvantages of increased weight, complexity, and moving parts.

With this simple bump detection method and the relatively small amount of data collected for this study, a Machine Learning approach was not utilized. For comparison in the future more training data will be collected so that a comparison to the current solution can be made.

Acknowledgement

This work was performed under the financial assistance award 70NANB17H173 from U.S. Department of Commerce, National Institute of Standards and Technology, PSCR Division and PSIA Program. This project is also funded in part by Carnegie Mellon University's Mobility21 National University Transportation Center, which is sponsored by the US Department of Transportation. The authors are grateful to the NIST PSCR Program Manager Jeb Benson for his comments and suggestions about the technical development of the hyper-reality helmet system.

References

[1] The thermal imaging drone Parrot website: www.parrot.com (2019)
[2] Hackett, S., Cai, Y. and Siegel, M., "Activity recognition from sensor fusion on fireman's helmet," Proceedings of CISP-BMEI, Suzhou, Oct. 19-21, (2020).
[3] Peabody, J. et al., Through-wall imaging radar, Lincoln Lab Journal, Vol. 19, No. 1, 2012
[4] Ultrasonic sensor range: https://www.quora.com/What-is-the-maximum-range-of-an-ultrasonic-sensor
[5] Structure from Motion, WikiPedia: https://en.wikipedia.org/wiki/Structure_from_motion
[6] Rowe, D. Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60(2), 91-110
[7] SURF: https://en.wikipedia.org/wiki/Speeded_up_robust_features
[8] FAST Corner detection: https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_feature2d/py_fast/py_fast.html
[9] SLAM: https://en.wikipedia.org/wiki/Simultaneous_localization_and_mapping
[10] Ramachandran, M. and Chellappa, R., Stabilization and mosaicing of airborne videos, International Conference on Image Processing, 2006
[11] Cai, Y., "Pheromone Trails", in Cai, Y. [Instinctive Computing], Springer-London, 2016
[12] Cai, Y. "Primitive Navigation", in Cai, Y., [Instinctive Computing], Springer-London, 2016
[13] Google Glasses: https://www.goggles4u.com/
[14] Hololens: https://www.microsoft.com/en-us/hololens
[15] Magic Leap: https://www.magicleap.com/
[16] Pokemon Game: https://www.headstuff.org/entertainment/gaming/the-future-of-ar-gaming-after-pokemon-go/

Author Biography

Yang Cai is founder of Visual Intelligence Studio, Senior Systems Scientist at CyLab Institute and Associate Research Professor of Biomedical Engineering at Carnegie Mellon University. He is the author of books: "Instinctive Computing" (Springer-London, 2016) and "Ambient Diagnostics" (CRC Press, 2014). His research area includes natural intelligence and augmented reality. His team won the NIST Haptic Interfaces for Public Safety Challenge in November 2019. Sean Hackett and Florian Alber are Research Engineers at Cylab and Ben Graham is Research Assistant and ECE student at Carnegie Mellon University. Mel Siegel is Professor of Robotics, Carnegie Mellon University. He is an IEEE Fellow and sensor and sensing technology expert.