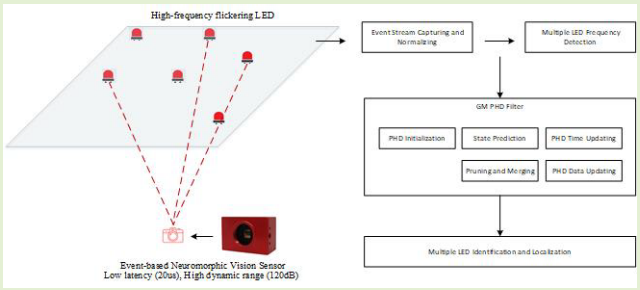
A Novel Visible Light Positioning System With Event-Based Neuromorphic Vision Sensor

Guang Chen, *Member, IEEE*, Wenkai Chen, Qianyi Yang, Zhongcong Xu, Longyu Yang, Jörg Conradt, *Senior Member, IEEE*, and Alois Knoll, *Senior Member, IEEE*

***Abstract*—With the advanced development of image processing technology, visible light positioning (VLP) system based on image sensors has attracted more and more attention. However, as a commonly used light receiver, traditional CMOS camera has limited dynamic range and high latency, which is susceptible to various lighting and environmental factors. Moreover, high computational cost from image processing is unavoidable for most of visible light positioning systems. In our work, a novel VLP system using an event-based neuromorphic vision sensor (event camera) as the light receiver is proposed. Due to the low latency and microsecond-level temporal resolution of the event camera, our VLP system is able to identify multiple high-frequency flickering LEDs in asynchronous events simultaneously leaving out the need for data association and traditional image processing methods. A multi-LED fusion method is applied and a high positioning accuracy of 3cm is achieved when the height between LEDs and the event camera is within 1m.**



***Index Terms*—Indoor localization, event camera, visible lighting positioning, GM-PHD filter.**

I. Introduction

WITH the development of internet-of-things (IoT) technology, connections between all kinds of domestic appliances and wireless devices are getting closer and closer. In the meantime, there is an increasing demand for indoor navigation and localization of intelligent devices, such as sweeping robot in house, service robot in canteen and automatic parking in underground parking lot. Thus, a great number of indoor location-based services (LBS) need to be implemented to improve human life [1], [2]. However, when applied to some complicated indoor surroundings like restaurants, school rooms or airport terminals, it is difficult to achieve a perfect LBS business that must guarantee a steady and high accurate positioning. In the past few years, lots of strategies about indoor localization have been proposed, which are mainly divided into two aspects: satellite positioning and indoor marker positioning. As for the satellite positioning, the agent carrying with satellite positioning system receives radio wave transmitted from several satellites orbiting around the earth in real time. Through measuring distances between the agent and these satellites, the agent's world coordinate about longitude and latitude could be computed directly [3], [4]. Up to now, different technologies of global navigation satellite system (GNSS) have been applied to many kinds of mobile devices and military equipment all around the world. Nonetheless, in the confined indoor space, the positioning accuracy of GNSS is not enough to satisfy the requirements for indoor robot's navigation due to the multi-path effect and radio disturbance. Hence, some researchers explore to install some access point (AP) devices inside the buildings, which could emit radio-frequency (RF) signal and their functions are the same as satellites [5], [6].

After obtaining the location information about these AP devices, the agent equipped with multiple signal receivers receives different RF signals and then its relative world position to the room is calculated. Although this method brings high localization accuracy, it is hard to guarantee information security due to the public signal protocols such as Wi-Fi, Bluetooth and Zigbee. Besides, several uncertain factors like

Manuscript received April 9, 2020; accepted April 12, 2020. Date of publication April 27, 2020; date of current version August 5, 2020. This work was supported in part by the European Union's Horizon 2020 Research and Innovation Program in Accordance under Grant 785907 (HBP SGA2), in part by the Shanghai Automotive Industry Sci-Tech Development Program under Grant 1838, and in part by the Shanghai AI Innovation Development Program 2018. The associate editor coordinating the review of this article and approving it for publication was Dr. Ioannis Raptis. *(Corresponding author: Guang Chen.)*

Guang Chen is with the School of Automotive Studies, Tongji University, Shanghai 201804, China, and also with the Chair of Robotics, Artificial Intelligence and Real-Time Systems, Technical University of Munich, 85748 Munich, Germany (e-mail: tj_autodrive@hotmail.com).

Wenkai Chen is with the School of Mechanical Engineering, Shanghai JiaoTong University, Shanghai 200240, China.

Qianyi Yang, Zhongcong Xu, and Longyu Yang are with the School of Automotive Studies, Tongji University, Shanghai 201804, China.

Jörg Conradt is with the KTH Royal Institute of Technology, 10044 Stockholm, Sweden.

Alois Knoll is with the Chair of Robotics, Artificial Intelligence and Real-Time Systems, Technical University of Munich, 85748 Munich, Germany.

Digital Object Identifier 10.1109/JSEN.2020.2990752

short transmitting distance, low-frequency capacity and change of signal intensity may hinder the continuous positioning.

As the visible light communication (VLC) and LED technology are getting more and more mature [7], [8], some researchers introduce visible LED light which is flickering at a high frequency into indoor localization. Without adding extra signal transmitting devices, visible light positioning (VLP) system can directly utilize LED luminaires installed in the room to achieve a high-accuracy positioning [9], [10]. LED luminaries are usually installed on the fixed region like ceiling, ensuring the stabilization and security for light emitting. Many traditional localization methods, such as Received Signal Strength (RSS), Time of Arrival (TOA), Angle of Arrival (AOA) and Time Difference of Arrival (TDOA) have been proposed to build a VLP system [11]. Focusing on color-based LEDs, indoor localization approaches by detecting the simultaneous transmission and subsequent transmission of RGB pulses are also proposed [12]. However, time synchronization and intensity attenuation require a high precision of visible light receivers. Besides, various LED luminaires have different shapes and their light diffusion couldn't be treated simply as the point diffusion, so the ranging error in traditional localization is inevitable. Recently, many researchers begin to use cameras to avoid the accuracy loss from traditional light receiver. When LED luminaires are flickering at a high frequency, it is difficult for ordinary CMOS camera to capture high-frequency features from images. Rolling shuffle effect is usually adopted to improve its dynamic range, with which many black fringe patterns could be recognized and analyzed through image processing. After that, by calculating average threshold value of grey-scale image or average linear intensity from fringe pattern vertically, different frequency number representing its relevant LED identification (ID) could be obtained [13], [14]. Nevertheless, flickering mitigation and dimming support are necessary to get a perfect image of LEDs. At the same time, when light intensity is strong enough or the agent moves in a high speed, blooming effect and motion blur phenomenon would disturb the rolling shuffle effect significantly. Though in [15], the authors have proposed a novel space-time RLS algorithm to equalize the motion blur from the optical channel, there is still a gap between the equalized results and natural high dynamic scenes due to the limit of hardware.

In addition, considering the limit of Field-Of-View (FOV) of camera, the number of LEDs in the image isn't constant while some tracking algorithms like particle filter should pre-determine parameter of targets number. Gaussian mixture probability hypothesis density (GM-PHD) filter is proposed to achieve multi-object tracking without knowing the number of targets [16]. Without the need of data association from the measurement and tracking, It is a modified recursion that propagates the first order statistic moment of random finite sets in time, to achieve objects tracking along with measurement uncertainty and negative recognition. Otherwise, the initial prior intensity and the posterior intensity of each random finite sets in GM-PHD filter are all regarded as Gaussian mixtures [17]. Because of the ability to solve variable targets and non-linear problem, it has been widely applied in the radar image [18], underwater scan [19] and human tracking [20]. [21] also proposed an improvement of GM-PHD filter based on weight penalization and multi-feature fusion, which could accurately track multi-objects with different features such as cells of different density and surveillance of different identities.

In this paper, we propose a practical multi-LEDs indoor localization system based on event camera [22] and GM-PHD filter tracking algorithm. Unlike CMOS cameras which has a fixed frame rate and limited dynamic range, the event camera naturally generates a stream of events with a microsecond time resolution and a high dynamic range (120dB), which is originally developed by [22]. Each pixel of an event camera triggers information independent of each other when that pixel detects a change of intensity. This information is called events that encode the precise triggering time, the pixel positions and sign of the intensity changes [22]–[26]. Without additional rolling shuffle effect, LEDs light flickering at a high frequency could be captured by the event camera precisely. Firstly, we normalize original event stream from event camera to remove background noise and low-frequency disturbance. Then, event histogram is established to detect all frequencies of multiple LEDs and each LED ID is determined referring to its frequency. A mapping table about LED ID and world position is made in advance. Thus, the world location of each LED could be obtained. Furthermore, no matter whether the agent is static or mobile, the GM-PHD filter module is adopted to track different LED luminaires in the image plane and each LED center coordinate could be extracted directly without complicated image processing, which ensures a stable and efficient localization. With the world coordinates and image coordinates of multiple LEDs, localization and median selection methods are applied to implement a high-accuracy positioning system after analysing the generalized imaging model of event camera. Several experiments show that the final positioning error is less than 3cm when the height between LED plane and event camera is within 1m.

II. The Proposed VLP System

A. System Architecture

As shown in Fig. 1, our proposed novel VLP system consists of a group of high-frequence flickering LEDs and an event-based neuromorphic vision sensor. Multiple LEDs are installed individually on the ceiling. The LEDs are flickering at a high frequency driven by the microcontroller. Each LED has a unique identification (L_i). Different from CMOS camera, event camera could capture illumination intensity changes at microsecond level while CMOS camera record the whole picture at a fixed frame rate (e.g., 20fps). The visualization of the output data of these two types of sensors is depicted in Fig. 2. It is clear to see that in a typical VLP system the event camera is a perfect light receiver that can capture the flickering of the LED (L_i) at a very high frequency, which is usually more than 200Hz that is far beyond the capability of traditional CMOS camera.

Considering various types and shapes of LED, it is important to represent the pixel coordinate of LED (I_i) on image plane. For traditional CMOS camera, a series of complex

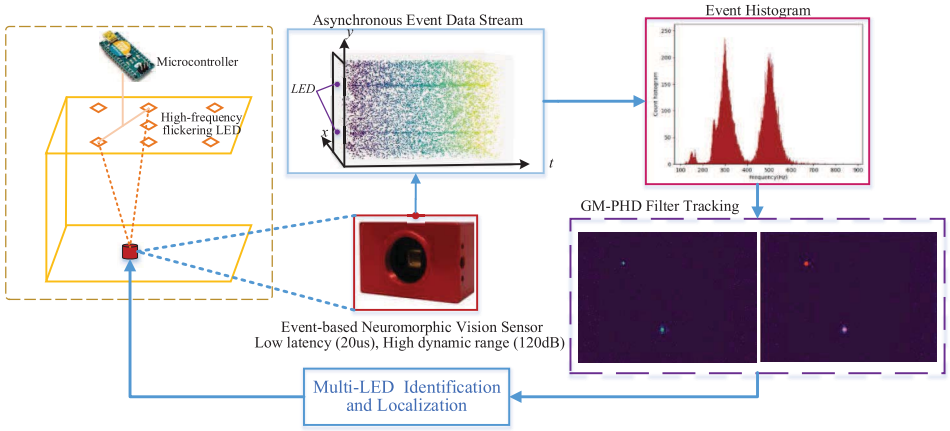


Fig. 1. Multi-LED VLP system architecture based on event camera.

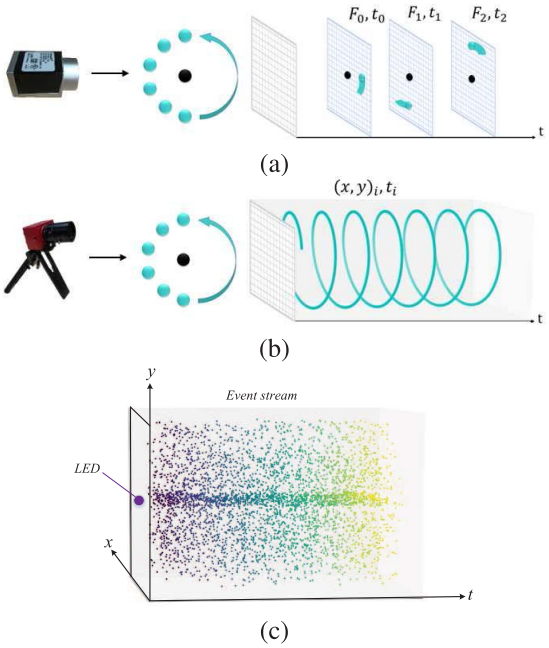


Fig. 2. (a) The CMOS camera captures all pixel intensities at a fixed frame rate. (b) The event camera captures intensity changes caused by the dynamic object. (c) Event stream from the event camera when a single LED is flickering.

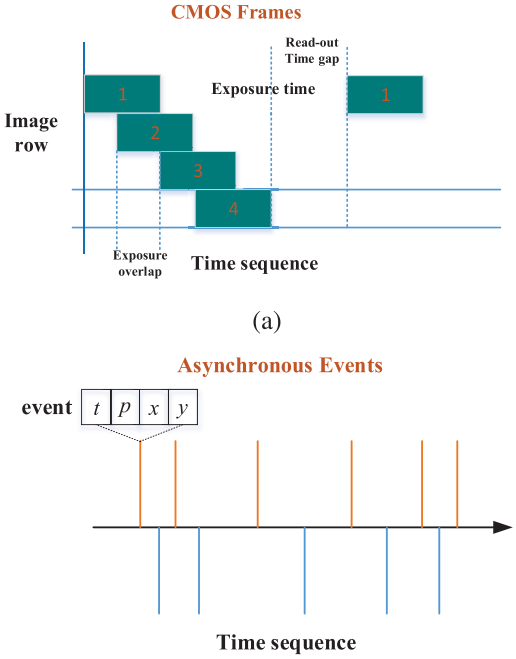


Fig. 3. Data reading mechanism. (a) CMOS frames. (b) Asynchronous event stream generated by event-based camera.

image processing steps are applied to extract LED center, such as gray scale, blur mask and threshold adjustment, but this method fails if the agent moves quickly or light intensity becomes strong. In our work, the LED center (X_i, Y_i) is extracted with GM-PHD filter algorithm, which ensures to achieve a steady localization in contrast with traditional VLP systems. In addition, the optical center of event camera in the world coordinate is expressed as $O(mw,nw,f_c)$, whose projection on the image plane is the image center and corresponding pixel coordinate is $O'(m,n)$. When the agent with event camera moves in different positions, geometry transformation and localization algorithm are applied to estimate agent's position. Finally, a multi-LED fusion strategy is used to calculate the optimal agent's location.

B. Event Camera Based LED Identification

In the process of visible light communication (VLC), modulation frequency of VLC is usually set to be greater than 200 Hz, it is a critical frequency which enables stable observation. In most VLP systems, rolling shuffle effect based on CMOS camera is implemented to detect such a high frequency flickering. Fig. 3 (a) shows detailed data reading mechanism of images recorded by CMOS camera, and it has some intrinsic drawbacks. Firstly, CMOS camera reads data row by row, which means that the activation of next row pixels has to wait for a full exposure of previous row. Hence, it is obvious that a time slice called exposure overlapping exists between each image row pixels. Secondly, when CMOS camera finishes a data frame, next image frame will be read until a time interval called read-out time gap ends. This time gap may cause a phenomenon called image artifact and a loss of communication data especially when camera moves. In addition, as the rolling shuffle effect, when light intensity is slightly high, blooming effect could cause the distortion of dark fringes and change the contrast ratio between bright fringes and dark fringes

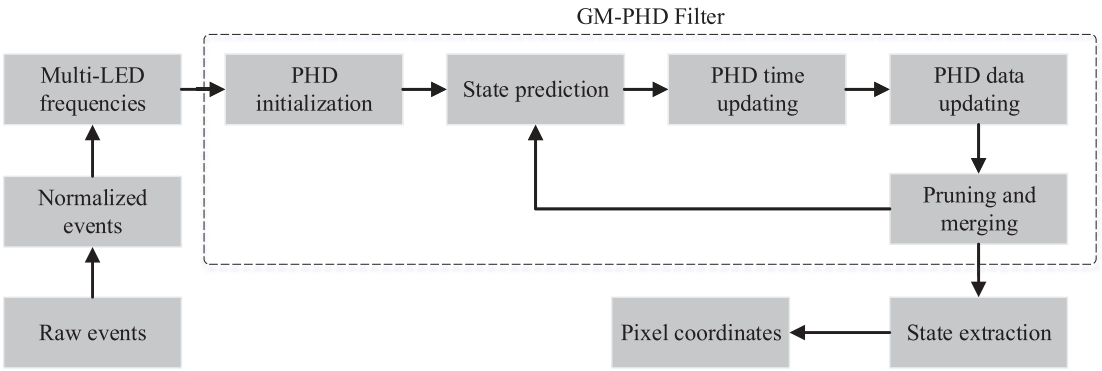


Fig. 4. Flow chart of Multi-LED tracking and extraction based on GM-PHD filter.

Fig. 5. Visualization of the event data captured by the event-based camera (a) original event (b) normalized event (c) histogram of frequencies calculated through intervals.

significantly which has side effects to the detection of certain LED frequency. Thus, these shortcomings of CMOS camera make it difficult to achieve a complete VLC and high-accuracy indoor positioning.

In our multi-LED positioning system, event-based camera is chosen as visible light receiver, which could generate a series of asynchronous event stream (See Fig. 3 (b)). The major difference between CMOS camera and event-based camera is that each pixel is independent in event-based camera. Each event consists of pixel coordinates of the event (x, y), polarity of the event p (positive or negative), and the timestamp t. In Fig. 3 (b), the upward and downward spikes represent positive and negative events respectively. By encoding only illumination changing direction, the bandwidth of transmitting, processing and storing this event stream is much lower than that of conventional cameras, which removes image redundancy significantly. Moreover, with a low latency $(20\mu s)$ and high dynamic range (120dB), the event camera is able to differentiate the high-frequency flickering LEDs and get rid of the distraction caused by the blooming effect and motion blur as the latency of event sensing pipeline is negligible compared to the dynamics of the environment.

1) LED Flickering Frequency Detection: Event camera, as the neuromorphic vision sensor, records the relative changes of light intensity, and each *event* can be described as $\{t_n, p_n, (x_n, y_n)\}$ where n is the index of event, $t_n(\mu s)$ is the time stamp of occurrence and (x_n, y_n) is the pixel coordinates of event in the field of view. When LED luminaires flicker at a high frequency, some pixels will appear as a positive event and a black speckle pattern could be observed in Fig. 5(a). To display the raw events clearly, we normalized event map in Fig. 5(b). In the progress of image normalizing, background motion and noisy data are imposed to a low value comparing with the events generated by blinking LED, which makes it obvious that LED is brighter than other pixels on the normalized event map.

Since the LED is flickering at certain frequency, polarity of events generated by LED will change in two types of transition, namely positive-to-negative or negative-to-positive with a timestamp t_{trans}. Inspired by [27], we calculate the intervals I_{trans} between the transitions with a same type which occur at the same pixel, the value of intervals nearly equals the flickering period of each LED luminaire, thus I_{trans} can serve as estimators of blinking frequency as $f = 1/I_{trans}$. Intervals can be described as $\{I_{trans}, (x, y), t\}$ where (x, y) is the pixel coordinate of transitions and t equals the timestamp t_{trans} of the first transition. As shown in Fig. 5(c), intervals accurately depict the blinking frequency from certain LED luminaires, the peak settles at the real blinking frequency of this luminaire, hence the certain LED can be detected and recognized, the frequency detection results can be found in section III. With the recognition of LED, a mapping table about LED ID and world position mentioned in Fig. 2 is referred to get the actual position of each LED light.

2) GM-PHD Filter Based Multi-LED Tracking: Most VLP systems based on CMOS camera use sophisticated image processing method to extract image center of LED. However, when light intensity of lights or background changes, it is necessary to reset threshold value to get target image. This method is complicated and inefficient especially when the agent is moving. Moreover, when the agent moves in the indoor environment, the LED number captured by the camera may change due to the limit of Filed-Of-View (FOV). Traditional object tracking methods like particle filter usually have to determine the number of target objects in advance, which obviously couldn't be applied in the VLP environment where the number of LEDs is not a priori. In this section, we present our GM-PHD filter based multi-LED tracking approach on the basis of the LED frequency detection, which is robust to the limit of FOV and cluttered environment variance.

Flow chart of GM-PHD tracker based on flickering events could be seen from Fig. 4. Firstly, raw event sequences are normalized to remove noise data and background disturbance. Then, normalized event map containing multi-LED frequencies information is input to a GM-PHD filter module. Detailed description about this tracker, especially designed for the flickering events, is clarified as follows:

Algorithm 1 GM-PHD Filter Algorithm With Event-Based Data

Input: w_t^l: GM components weights; μ_t^l: GM components means; Σ_t^l: GM components deviations; $C_t : \{c_1, c_2, \ldots, c_N\}$: GM components; t: Update time;
Output: pixel coordinates (x_i, y_i);
Initialize $F_{t|t}(s|M_t) = \sum_{l=1}^{L_t} w_t^l \mathcal{N}(s; \mu_t^l, \Sigma_t^l)$;
for *each timestep* t **do**
  Feature extraction: event transition intervals I_{trans};
  GMPHD_StatePredictionForBirthTargets;
  for $l \in L_t$ **do**
    GMPHD_StatePredictionForExistTargets;
    GMPHD_TimeUpdating;
    GMPHD_DataUpdating;
  PruningAndMerging;

(1) When time $t = 0$, the PHD $F_{t|t}(s|M_t)$ of whole event map is initialized with a weighted sum of Gaussian items: $F_{t|t}(s|M_t) = \sum_{l=1}^{L_t} w_t^l \mathcal{N}(s; \mu_t^l, \Sigma_t^l)$. As a Bayes probabilistic filter, M_t means the sets of multi-LED measurements, s means the corresponding multi-LED states. They are all distributed across the state space where each Gaussian item consists of a weight w_t^l, mean value μ_t^l and variance Σ_t^l. And L_t is the pre-defined number of Gaussian components in the PHD filter.

(2) When time $t \geq 1$, the predicted PHD after initializing and time updating is regarded as a Gaussian mixture, which could be expressed as: $F_{t|t-1}(s) = F_{G,t|t-1}(s) + \gamma_k(s)$. Among this equation, $F_{G,t|t-1}(s)$ represents the existing LEDs status in the FOV while $\gamma_k(s)$ means the status of new-coming or missing LEDs caused by the variation of FOV, where the change of Gaussian intensity could be found. Their equations could also be expressed as follows:

$$F_{G,t|t-1}(s) = Ps \sum_{l=1}^{L_{t-1}} w_{t-1}^l \mathcal{N}(s; \mu_{G,t|t-1}^l, \Sigma_{G,t|t-1}^l)$$
$$\gamma_k(s) = \sum_{l=1}^{L_{r,k}} w_{r,k}^l \mathcal{N}(s; \mu_{r,k}^l, \Sigma_{r,k}^l)$$

Otherwise, as the PHD data updating, a Gaussian mixture also could approximate the updated PHD measurements: $F_{t|t}(s) = \sum_{l=1}^{L_t} w_{t|t}^l \mathcal{N}(s; \mu_t^l, \Sigma_t^l)$. Detailed GM-PHD filter time updating and data updating algorithm could be seen in [16].

(3) Due to the change of LEDs number in the FOV, it is much important to prune and merge updated Gaussian components. When the weights of a Gaussian component is below a pre-defined threshold w_p, the updated PHD $F_{t|t}(s|M_t)$ will be removed. Similarly, when the mean distance of different Gaussian components is below a pre-determined merging threshold w_m then the updated PHD $F_{t|t}(s|M_t)$ also will be merged. The strategy of pruning and merging could eliminate some disqualified Gaussian components, which could reduce the computational complexity significantly.

Algorithm 1 specifies the procedure of GM-PHD filter on the multi-LED event data. It is noted that whole PHD filter module is updated, pruned, merged and estimated to match

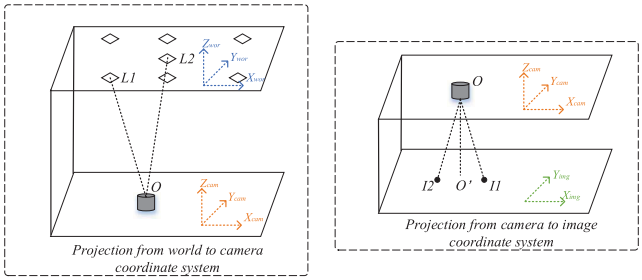


Fig. 6. System geometry transformation from world to image coordinate.

multi-LED frequencies using transition intervals I_{trans}, which are obtained from two successive transitions of the same type. Seen from Fig. 5(c), the reciprocal of transition interval could be approximated as a Gaussian distribution, where standard deviation is corresponding with the peak value. The transition intervals can serve as estimators of multi-LED frequencies and different intervals at pixel coordinates (x, y) also convey the probability that an LED which is flickering at a frequency of $1/I_{trans}$ is detected at (x, y), guiding the Gaussian components of GM-PHD filter to update. The weights $w_{t|t}^l$ are calculated based on $1/I_{trans}$ according to the closed PHD update algorithm and the mean $\mu_{t|t}^l$ and covariance $\Sigma_{t|t}^l$ are updated with the Kalman filter update algorithm. Finally, after pruning and merging, pixel coordinates (x_i, y_i) and their Gaussians G_i could be obtained through calculating weighted filter output. In our system, due to limited experimental number of LED, the maximum number of Gaussian mixture components is set as 30. And the pre-determined threshold of pruning and merging are set to 10^{-15} and 5 separately.

C. Multi-LED Position Estimation

In our system, 6 LED luminaires are installed on the ceiling and their status (ON/OFF) are controlled by the LED driver. And the LED driver outputs different PWM signals to activated LED at a certain frequency number and duty cycle. As shown in Fig. 5, system geometry transformation is described to calculate camera's position by combining world positions of LEDs with corresponding pixel coordinates. Firstly, to reduce model computing complexity, two LED luminaires L1, L2 are flickering on the ceiling and their lights are received by the event camera successfully, which achieves the projection from world coordinate to camera coordinate. The world position of L1 and L2 are expressed as (XW_1, YW_1, Z_w), (XW_2, YW_2, Z_w). The value Z_w represents the distance from ceiling plane to event camera's optical center. Secondly, according to the principle of pinhole imaging and event camera's optical attribute, L1 and L2's image coordinates are projected from the camera coordinate. It is noted that our system regards the z-direction coordinate value of image plane as zero. So, two LEDs' image center and pixel coordinates are expressed as $I_1(X_1, Y_1, 0)$, $I_2(X_2, Y_2, 0)$. Similar to CMOS camera, the focal length f_c represents the vertical distance between optical center and image plane. The optical center and its projection on image plane are expressed as $O(mw, nw, f_c)$ and $O'(m, n)$. Then, the theory of similar triangles is applied to solve camera's world coordinates

$O(mw, nw, f_c)$ based on the pixel coordinates of different LED lights. Detailed calculation procedure is expressed as follows: Firstly, the real distance Wo_dist between L1 and L2 and pixel distance Pi_dist between $I1$ and $I2$ could be acquired:

$$Wo\_dist = \sqrt{(XW_2 - XW_1)^2 + (YW_2 - YW_1)^2} \quad (1)$$

$$Pi\_dist = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (2)$$

$$k = \frac{Wo\_dist}{Pi\_dist} \quad (3)$$

$$Z_{diff} = k \times f_c \quad (4)$$

After finishing the calibration of event camera and obtaining the ratio k from Wo_dist and Pi_dist, vertical distance Z_{diff} between the event camera and LED luminaires could be derived using f_c directly. Available from equations (1) – (3) and image center coordinate $O'(m, n)$, then the camera's world position about xy-axis (mw, nw) could be estimated:

$$O_{xy} = \begin{pmatrix} mw \\ nw \end{pmatrix}, \quad L1_{xy} = \begin{pmatrix} XW_1 \\ YW_1 \end{pmatrix}, \quad L2_{xy} = \begin{pmatrix} XW_2 \\ YW_2 \end{pmatrix} \quad (5)$$

$$A = \begin{pmatrix} X_1 - m \\ Y_1 - n \end{pmatrix}, \quad B = \begin{pmatrix} X_2 - m \\ Y_2 - n \end{pmatrix} \quad (6)$$

$$O_{xy} = k \times A + L1_{xy}, \quad or \ O_{xy} = k \times B + L2_{xy} \quad (7)$$

where A is the pixel distance between LED lighting image $I1$ and image center O', B is the pixel distance between the other LED lighting image $I2$ and image center O'. O_{xy} is calculated as the final predicted position of camera.

In practice, the ceiling is not absolute horizontal that multiple LED luminaires couldn't be installed so flat, which leads to an unavoidable location error [28]. Hence, similar geometric relationship between multi-LED pairs could be applied to calculate estimated position of event camera separately. For example, the LED number in the image plane is N $(N \geq 2)$, then possible number of camera position value is C_N^2. A strategy of arithmetic mean is proposed to use all potential camera position value to solve this misalignment [29]. However, when multiple LEDs are captured in our image plane, some position outliers may be calculated once certain LED ID couldn't be recognized correctly due to the change of environmental or signal conditions. It is obvious that arithmetic mean couldn't solve it properly. Hence, we propose a median selection method to find the optimal camera position. After acquiring all potential camera positions O_{xyi} (mw_i, nw_i) $(i = 1, 2 \ldots C_N^2)$, all x-axis and y-axis coordinate values mw_i, nw_i are arranged in an ascending order separately. Final estimated position could be expressed as O_{xyf} $\left(mw_f, nw_f\right)$ and the median selection method is used:

$$mw_f = \frac{mw_{C_N^2/2} + mw_{C_N^2/2+1}}{2},$$

$$nw_f = \frac{nw_{C_N^2/2} + nw_{C_N^2/2+1}}{2} \quad (8)$$

$$mw_f = mw_{(C_N^2+1)/2}, \quad nw_f = nw_{(C_N^2+1)/2} \quad (9)$$

where $mw_{C_N^2/2}$ and $mw_{C_N^2/2+1}$ represents the x-axis coordinate value at the arranged sequence of $(C_N^2/2)^{th}$ and $((C_N^2+2)/2)^{th}$. If C_N^2 is an even number, equation (8) is used to get final camera position. Otherwise, if C_N^2 is an odd number, final camera position could be computed from equation (9).

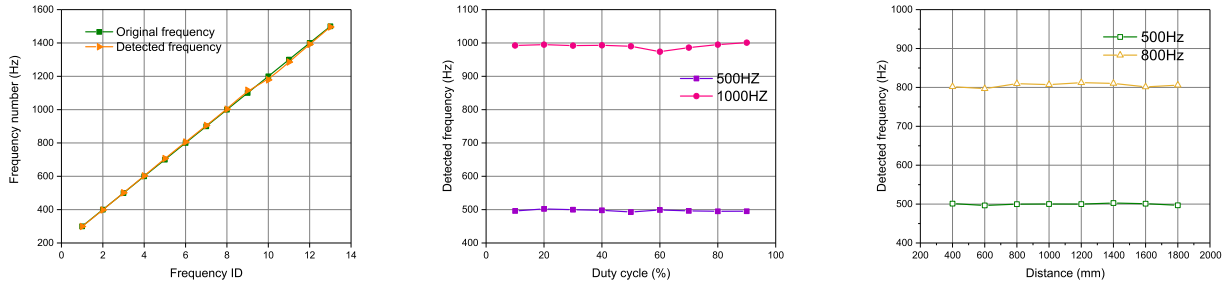
III. Experiment Setup and Results

In order to evaluate the effects of different multi-LED combinations, six LED luminaires are respectively installed at (200, 100), (500, 100), (200, 500), (350, 500), (500, 500), (350,300), where the coordinate unit is mm. We use Arduino Nano V3.0 micro-controller boards to control their on-off status. Different PWM signals are transmitted to the specific port of the controller, which ensures that each LED is flickering at a high frequency and assigned with a unique ID. For instance, L1, L2, and L3 could be regarded as a multi-LED combination and they are given the frequencies of 500 Hz, 400 Hz, and 300 Hz with a same duty ratio of 50%. As stated in Section 2, frequency number, the unique ID and corresponding world coordinate of each LED luminaires are also pre-stored in a mapping table. In addition, we adopt Davis 346Red as the event camera module in all experiments. As the most advanced event camera, it has a resolution of 340 × 246 pixel, a dynamic range up to 120dB and its minimum delay could be as low as $20\mu s$. It also could be directly connected to the computer through USB 3.0 terminal and then dynamic events are observed and recorded through the open source software jAER. To evaluate the proposed VLP system, we set 15 test points (TP_s) where event camera could capture and detect all multi-LED combinations. These test points are divided into 3 rows and 5 columns and each row distance is 50mm, each column distance is 100mm. The coordinate of first test point TP_1 is determined at (250,200). Other test points could be calculated using module operation, whose coordinates are represented as $TP_i = (200 + 50(i-1), 100(2 + i\%5)), i = 2, 3, 4, 5$ and $TP_i = (200 + 50(i\%5 - 1), 100(2 + i\%5)), i = 6, 7, \ldots, 15$.

A. Frequency Detection Results

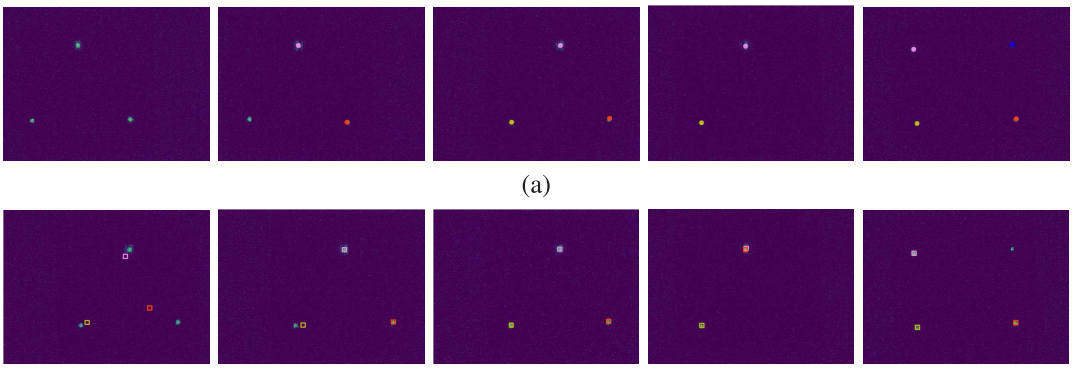
Accurate frequency detection is vital to affect our system localization performance, especially there exists multiple LED luminaires in the image plane. Once a LED's frequency isn't recognized correctly, whose localization information from the mapping table will be recognized wrong obviously. In this part, L6 is activated by the PWM signal in a series of frequencies from 300Hz to 1500Hz and duty cycles from 10% to 90%. Besides, the height between LED luminaries and event camera varies from 400mm to 2000mm.

As Fig. 7(a) shows, the event camera could detect the flickering frequency stably when duty cycle is set at 50% and original height is fixed at 1000mm. With the flickering frequency increasing, accurate frequency detection demonstrates that a wide dynamic range from the event camera is practical without rolling shuffle effect. Otherwise, when flickering frequency is constant, Fig. 7(b)(c) shows the impact of potential factors on frequency detection, such as duty cycle and limited distance between the camera and LEDs. As we can see, when duty cycle is higher or limited distance is farther, the event camera is still able to accurately identify different frequencies. Hence,



(a) (b) (c)

Fig. 7. The difference between original frequency and detected frequency. (a) shows a range of frequencies from 300Hz to 1500Hz detected from a fixed distance of 1m (b) shows two different frequencies, 500Hz and 1000Hz, detected through the transition of duty cycle from 10% to 90% (c) shows two frequencies, 500Hz and 800Hz, detected by changing the distance from 400mm to 2000mm.



(a)

(b)

Fig. 8. Visualized tracking process for LED events: LEDs flickering in 300Hz, 400Hz, 500Hz and 600Hz (green), different tracking targets (yellow, red, violet, blue) and (a) GM-PHD filter tracker (b) particle filter tracker.

it effectively improves the shortcomings of traditional CMOS camera that is susceptible to extra environmental factors [13].

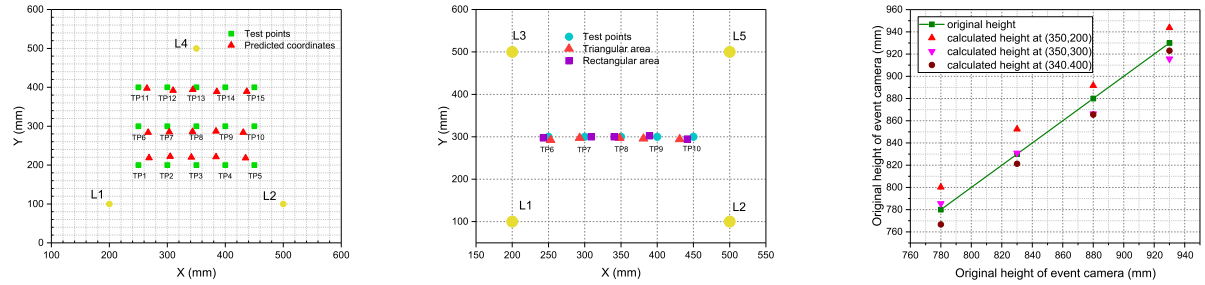
B. Tracking and Extraction Results

To achieve continuous VLC and indoor positioning, it is also important to track and extract multi-LED image centers when the agent is at a motion. And it is much difficult and costly for traditional image processing method to extract pixel center. In this part, GM-PHD filter algorithm is adopted to calculate multi-LED pixel coordinates successfully without pre-determining the number of LEDs. Three LED luminaries, L1, L2, L4 are firstly selected to flicker in a certain frequency such as 300Hz, 400Hz, 500Hz. To simulate the newborn and missing targets from the FOV, two processes are conducted successively: L2 is firstly removed from the system then L2, L3 and L5, flickering at 400Hz, 500Hz and 600HZ, are added to the system. With these different frequency data as target values, the GM-PHD filter is used for tracking multi-LED event centers. Fig. 8 shows the tracking process and comparison results for these LED events from the GM-PHD filter tracker and traditional particle filter tracker.

When the number of LEDs captured by event camera decreases at a specific FOV, the phenomenon of particles overlapping and identification error emerges in the particle filter tracker. More importantly, a fixed particle number of the tracker obviously couldn't recognize and track newborn LEDs. As our GM-PHD filter tracker, it can be seen that L1, L2, L3, L4 and L5 are all tracked successfully. After finishing the tracking, all pixel coordinates of multiple LEDs are extracted through calculating the weighted filter output. In our experiment, when L1, L2, L4 are flickering, the event camera moves laterally from TP_6 to TP_{10} and then error analysis is done by comparing our extracting results with ground truth of manual annotation at each TP. The Root Mean Square Error (e_x,e_y) and Standard Deviation (σ_y) statistics are shown in Table I. Furthermore, inspired by the work of [30], [31] in the LED-camera VLP system, the Cramer-Rao lower Bound (CRLB) results of pixel accuracy in the visualized event image are also computed as a benchmark to evaluate our estimation performance. Due to the unknown measurement noise and signal noise from event-based camera [32], each LED is adapted to a point spread function (PSF) to approximate its intensity distribution roughly and the LED image model is expressed as: $I_{x,y} = A\exp(-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}) + B$, where A, σ and B represent the peak intensity, PSF width and background intensity, x_0 and y_0 are the ground truth of image center. If not considering the experimental equipment jitter and measurement error, we find that our method using GM-PHD filter to extract image center is practical.

C. Event Camera Based Positioning Performance

After completing the experiments in the above two parts, the world position of event camera could be estimated using proposed multi-LED positioning method. Considering the



(a) (b) (c)

Fig. 9. Localization performance of the proposed multi-LED positioning system (a) shows a stable positioning error within 3cm at all test points (b) shows an accurate track under different multi-LED combinations (c) describes the height estimation from event camera to LED luminaires.

TABLE I
RMSE (PIXEL), SD (PIXEL) AND CRLB (PIXEL) OF PIXEL POSITION ESTIMATE AT DIFFERENT FLICKERING FREQUENCIES

Flickering frequency (Hz)	e_x	e_y	σ_y	$CRLB$
300	0.975	2.316	0.159	0.156
400	3.269	3.547	0.688	0.518
500	4.196	4.328	0.464	0.561

difference of agent's position and multi-LED combinations, three experiments are done to validate final positioning accuracy in all dimensions. In the first experiment, fixed at a vertical distance of 1000mm, L1, L2 and L4 are activated at a duty cycle of 50% and frequency of 300Hz, 400Hz and 500Hz.

Fig. 9(a) shows a stable positioning error within 3cm between the test points (green marker) and estimated coordinates (red marker). In the second experiment, the agent moves from TP_6 to TP_{10} and LED luminaires with different flickering frequencies are fixed at two combinations—triangular (L1, L2, L3) and rectangular (L1, L2, L3, L5). Same to the first experiment, the distance between event camera and multi-LED plane is set at 1000mm. It can be seen in Fig. 9(b) that different LED combinations does not change much as the localization accuracy is always maintained. In the last experiment, the focus length of event camera is applied to calculate vertical distance from event camera to LED luminaires. With original height increasing from 780mm to 930mm, estimated heights at different $TP_s(TP_3, TP_8, TP_{13})$ are depicted in Fig. 9(c). It could be seen that the measurement errors are also limited in 3 cm at different heights.

IV. CONCLUSION

This work implements a novel indoor localization system by integrating a new LED light receiver, *the event-based neuromorphic vision sensor (event camera)*, into a VLP system. Conventional VLP system often uses the CMOS camera as the receiver that relies on traditional image processing methods, which consume too much computational resource. In contrast, due to the low latency and microsecond-level temporal resolution of the event camera, our VLP system is able to track multiple high-frequency flickering LED simultaneously without a need of data association and traditional image processing methods. The GM-PHD filter is applied to identify LED flickering frequency and positions. After the fusion of multi-LED position estimations, final localization results reveal that a high positioning accuracy of 3cm could be achieved when the height between LED plane and event camera is within 1m.

REFERENCES

[1] P. Davidson and R. Piche, "A survey of selected indoor positioning methods for smartphones," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1347–1370, 2nd Quart., 2017.
[2] J. Luo, L. Fan, and H. Li, "Indoor positioning systems based on visible light communication: State of the art," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2871–2893, Aug. 2017.
[3] Y. Lei, T. Wang, and J. Wu, "Vehicles relative positioning based on ZigBee and GPS technology," in *Proc. 6th Int. Conf. Electron. Inf. Emergency Commun. (ICEIEC)*, Jun. 2016, pp. 59–62.
[4] M. Yoshino, S. Haruyama, and M. Nakagawa, "High-accuracy positioning system using visible LED lights and image sensor," in *Proc. IEEE Radio Wireless Symp.*, Jan. 2008, pp. 439–442.
[5] W.-H. Kuo and H.-H. Shih, "A novel AP positioning scheme under information-limited indoor environments," in *Proc. IEEE 6th Global Conf. Consum. Electron. (GCCE)*, Oct. 2017, pp. 1–2.
[6] H.-X. Zhao and J.-T. Wang, "A novel three-dimensional algorithm based on practical indoor visible light positioning," *IEEE Photon. J.*, vol. 11, no. 3, pp. 1–8, Jun. 2019.
[7] N. Chi, H. Haas, M. Kavehrad, T. D. C. Little, and X.-L. Huang, "Visible light communications: Demand factors, benefits and opportunities [Guest Editorial]," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 5–7, Apr. 2015.
[8] X. Huang *et al.*, "2.0-Gb/s visible light link based on adaptive bit allocation OFDM of a single phosphorescent white LED," *IEEE Photon. J.*, vol. 7, no. 5, pp. 1–8, Oct. 2015.
[9] M. F. Keskin, A. D. Sezer, and S. Gezici, "Localization via visible light systems," *Proc. IEEE*, vol. 106, no. 6, pp. 1063–1088, Jun. 2018.
[10] J. Armstrong, Y. Sekercioglu, and A. Neild, "Visible light positioning: A roadmap for international standardization," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 68–73, Dec. 2013.
[11] T. Akiyama, M. Sugimoto, and H. Hashizume, "Time-of-arrival-based smartphone localization using visible light communication," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2017, pp. 1–7.
[12] S. Pergoloni, Z. Mohamadi, A. M. Vegni, Z. Ghassemlooy, and M. Biagi, "Metameric indoor localization schemes using visible lights," *J. Lightw. Technol.*, vol. 35, no. 14, pp. 2933–2942, Jul. 15, 2017.
[13] Y. Han, Q. Cheng, and P. Liu, "Indoor positioning based on LED-camera communication," in *Proc. IEEE Int. Conf. Consum. Electron.-China (ICCE-China)*, Dec. 2016, pp. 1–4.
[14] J. Fang *et al.*, "High-speed indoor navigation system based on visible light and mobile phone," *IEEE Photon. J.*, vol. 9, no. 2, pp. 1–11, Apr. 2017.
[15] S. Pergoloni, M. Biagi, S. Colonnese, R. Cusani, and G. Scarano, "A space-time RLS algorithm for adaptive equalization: The camera communication case," *J. Lightw. Technol.*, vol. 35, no. 10, pp. 1811–1820, May 15, 2017.
[16] B. N. Vo and W. K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.

[17] M. Yazdian-Dehkordi, Z. Azimifar, and M. A. Masnadi-Shirazi, "Competitive Gaussian mixture probability hypothesis density filter for multiple target tracking in the presence of ambiguity and occlusion," *IET Radar, Sonar Navigat.*, vol. 6, no. 4, pp. 251–262, 2012.
[18] M. Tobias and A. D. Lanterman, "Probability hypothesis density-based multitarget tracking with bistatic range and Doppler observations," *IEE Proc.-Radar, Sonar Navigat.*, vol. 152, no. 3, pp. 195–205, 2005.
[19] D. Clark and J. Bell, "Multi-target state estimation and track continuity for the particle PHD filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 99, pp. 1441–1453, Oct. 2007.
[20] Z. Fu, F. Angelini, S. M. Naqvi, and J. A. Chambers, "GM-PHD filter based online multiple human tracking using deep discriminative correlation matching," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 4299–4303.
[21] X. Zhou, H. Yu, H. Liu, and Y. Li, "Tracking multiple video targets with an improved GM-PHD tracker," *Sensors*, vol. 15, no. 12, pp. 30240–30260, Dec. 2015.
[22] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128 × 128 120 dB 15 μs latency asynchronous temporal contrast vision sensor," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, 2008.
[23] J. Conradt, R. Berner, M. Cook, and T. Delbruck, "An embedded AER dynamic vision sensor for low-latency pole balancing," in *Proc. IEEE 12th Int. Conf. Comput. Vis. Workshops (ICCV Workshops)*, Sep. 2009, pp. 780–785.
[24] D. Weikersdorfer and J. Conradt, "Event-based particle filtering for robot self-localization," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2012, pp. 866–870.
[25] G. Chen, H. Cao, J. Conradt, H. Tang, F. Röhrbein, and A. Knoll, "Event-based neuromorphic vision for autonomous driving: A paradigm shift for bio-inspired visual sensing and perception," *IEEE Signal Process. Mag.*, 2020, doi: 10.1109/MSP.2020.2985815.
[26] G. Chen, L. Hong, J. Dong, P. Liu, J. Conradt, and A. Knoll, "EDDD: Event-based drowsiness driving detection through facial motion analysis with neuromorphic vision sensor," *IEEE Sensors J.*, early access, Feb. 10, 2020, doi: 10.1109/JSEN.2020.2973049.
[27] A. Censi, J. Strubel, C. Brandli, T. Delbruck, and D. Scaramuzza, "Low-latency localization by active LED markers tracking using a dynamic vision sensor," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 891–898.
[28] J.-Y. Kim, S.-K. Han, S.-H. Yang, and Y.-H. Son, "High-resolution indoor positioning using light emitting diode visible light and camera image sensor," *IET Optoelectron.*, vol. 10, no. 5, pp. 184–192, Oct. 2016.
[29] J.-K. Lain, L.-C. Chen, and S.-C. Lin, "Indoor localization using K-Pairwise light emitting diode image-sensor-based visible light positioning," *IEEE Photon. J.*, vol. 10, no. 6, pp. 1–9, Dec. 2018.
[30] X. Zhao and J. Lin, "Theoretical limits analysis of indoor positioning system using visible light and image sensor," *ETRI J.*, vol. 38, no. 3, pp. 560–567, Nov. 2015.
[31] E. Curry. (2016). *Parameter Estimation in an Led-Camera Visible Light Communication System*. [Online]. Available: http://hdl.handle.net/10150/624188
[32] C. Brandli, R. Berner, M. Yang, S.-C. Liu, and T. Delbruck, "A 240 × 180 130 dB 3 μs latency global shutter spatiotemporal vision sensor," *IEEE J. Solid-State Circuits*, vol. 49, no. 10, pp. 2333–2341, Oct. 2014.

Guang Chen (Member, IEEE) received the B.S. and M.Eng. degrees in mechanical engineering from Hunan University, China, and the Ph.D. degree from the Faculty of Informatics, Technical University of Munich, Germany. He is a Research Professor with Tongji University and a Senior Research Associate (Guest) with the Technical University of Munich. He is leading the Intelligent Perception and Computing System Group, Tongji University. His research interests include computer vision, image processing and machine learning, and the bio-inspired vision with applications in robotics and autonomous vehicle. He was awarded the Program of Tongji Hundred Talent Research Professor in 2018.

Wenkai Chen received the B.S. degree in mechanical engineering from Wuhan University, Wuhan, China, in 2017. He is currently pursuing the M.S. degree in mechanical engineering with Shanghai Jiao Tong University, Shanghai, China. From 2018 to 2019, he worked as a Research Assistant with the Intelligent Perception and Computing Group, Tongji University.

Qianyi Yang is currently pursuing the B.E. degree in vehicle engineering with Tongji University, Shanghai, China. From 2018 to 2019, he worked as a Research Assistant with the Intelligent Perception and Computing Group, Tongji University. His research interests include the autonomous vehicle and UWB wireless positioning.

Zhongcong Xu is currently pursuing the B.E. degree in vehicle engineering with Tongji University, Shanghai, China. Since 2018, he has been working as a Research Assistant with the Intelligent Perception and Computing Group, Tongji University. His research interests include autonomous vehicles and machine vision.

Longyu Yang is currently pursuing the B.E. degree in computer science and technology with Tongji University, Shanghai, China. Since 2019, he has been working as a Research Assistant with the Intelligent Perception and Computing Group, Tongji University. His research interest lies in the intersection part of computer vision and natural language processing, including image captioning, visual question answering, and video captioning.

Jörg Conradt (Senior Member, IEEE) received the Diploma degree in computer engineering from TU Berlin, the M.S. degree in robotics from the University of Southern California, and the Ph.D. degree from ETH Zürich. He is an Associate Professor with the School of Electrical Engineering and Computer Science, KTH, Stockholm, Sweden. Before joining KTH, he was a W1 Professor at TUM, Germany. He was the Founding Director of the Elite Master Program NeuroEngineering at TUM. Research in his group on Neurocomputing Systems investigates key principles by which information processing in brains works, and applies those to real-world interacting technical systems.

Alois Knoll (Senior Member, IEEE) received the Diploma (M.Sc.) degree in electrical/communications engineering from the University of Stuttgart, Germany, in 1985, and the Ph.D. *(summa cum laude)* in computer science from the Technical University of Berlin, Germany, in 1988. He served on the faculty of the Computer Science Department, TU Berlin, in 1993. He joined the University of Bielefeld, as a Full Professor and the Director of the Research Group Technical Informatics in 2001. Since 2001, he has been a Professor with the Department of Informatics, TU München. From 2004 to 2006, he was the Executive Director of the Institute of Computer Science, TUM. He was the Program Chairman of the IEEE Humanoids2000, the General Chair of the IEEE Humanoids2003, the Program Chair of the IEEE-IROS 2015, and the Editor-in-Chief of *Frontiers in Neurorobotics*.