

LOSA-X: Expandable 3D Motion Tracking System

Hailin Ren, *Student Member, IEEE*, Anil Kumar, *Member, IEEE*, and Pinhas Ben-Tzvi¹, *Senior Member, IEEE*

Abstract—This paper presents a novel, low-cost, expandable 3D motion tracking system with the ability to adapt to dynamic background illumination. The proposed system consists of multiple tracker modules equipped with a linear optical sensor arrays and a 9-DOF inertial measurement unit for position and orientation tracking of a wireless IR-LED illuminated active marker. This paper presents an extended Kalman filter-based sensor fusion architecture for combining data acquired from multiple trackers and marker for enhanced accuracy and operational range. The marker system adaptively adjusts the LED intensity to optimize energy consumption and measurement accuracy. The proposed system's ability to extract background illumination allows it to be robust toward differences in scene illuminations, both static and dynamic. Detailed hardware design and performance evaluation through experimentation have been discussed in this paper.

Index Terms—3D motion tracking, expandable workspace, background removal, adaptive sensing.

I. INTRODUCTION

THE ability to track three-dimensional motion is vital in many industrial and academic research applications including machine vision and robotics. Applications such as robot navigation [1], [2], virtual reality [3], [4] and augmented reality [5], [6] require real time processing and low-latency motion capture systems. Medical applications such as gait analysis [7], [8], filmmaking, and video development applications require systems with high accuracy to reconstruct the 3D pose of actors [9] or capture subtle movements such as facial expressions [10]. Applications like virtual reality and augmented reality heavily rely on accurate motion tracking while recording the subject in real-time for the construction of a virtual scenario. This requirement is not limited to applications in structured indoor environments. Motion capture systems are also deployed for outdoor applications [11] and underwater activities [12].

Depending on the type of object being tracked, motion capture systems can be classified into two categories: marker-based [13] and marker-less systems [14]. Marker-less systems do not require subjects to wear special equipment for tracking. Usually, such systems rely on imagery obtained from multiple cameras and track the 3D representation of the subject using

optical flow based methods [15]. To improve accuracy, such systems opt for a model based approach to extract and track parts of interest [16].

In contrast to marker-less, marker-based motion capture systems are often self-contained systems and require less post-processing for accurate tracking. Such systems often rely on sensing modalities like, radio wave [17], inertial [18], LED [19], magnetic, retro-reflective [20], or markers of any combination therein to track motion [21]. Based on the marker's illumination, marker-based systems can be divided into passive markers [22], activate markers, or semi-passive marker systems [23].

Passive marker systems deploy simple and lightweight reflective markers onto subjects to limit the effects of the marker on the motion of subject. Most commercial motion tracking systems like Optitrack™ [24], Vicon™ [25], and Qualisys™ [26] use reflective markers which are coated with a retroreflective material to reflect light that is generated near the camera lenses. These kinds of systems require multiple cameras to be located around the subjects to cover the entire motion. Active marker systems such as Phasespace™ [27] utilize different frequency LED lights to generate unique identifiers for each LED, which also solves the correspondence issue. Active markers are designed to emit light by themselves and the tracking system triangulates their positions. Multiple markers with different color LEDs can be used to track specific subjects. Such marker systems usually have a high signal to noise ratio and high measurement resolution which can degrade with distance (inverse square law). Another kind of marker system is the semi-passive imperceptible marker system, such as in Prakash [23]. Prakash uses an LED with a passive binary mask as the transmitter and a photosensor as the receiver. A set of optical transmitters are placed to encode the volume of interest spatiotemporally while photosensors demultiplex the coded optical signals from multiple transmitters to compute the receiver's position and orientation. Besides the optical marker system mentioned above, other marker systems such as Pozyx™ [17] utilize ultra-wideband RF technology for positioning. For 3D positioning, at least five modules are needed, where four non-coplanar modules work as anchors and track position of another module installed on the subject. Many researchers have explored use of line scan cameras based stereovision systems for estimating motion of a LED light source (marker) [28]–[30]. However such systems, while using expensive sensors and optics, rely only on vision sensing which have limited field of view and suffer significant performance degradation with increased distance.

Whereas commercial motion tracking systems such as Vicon, Phasespace and Optitrack are too expensive for general use, cheaper systems such as Pozyx and Kinect [31] are not

Manuscript received November 7, 2018; revised December 20, 2018; accepted December 28, 2018. Date of publication January 21, 2019; date of current version April 5, 2019. The associate editor coordinating the review of this paper and approving it for publication was Prof. Giancarlo Fortino. (Hailin Ren and Anil Kumar contributed equally to this work.) (Corresponding author: Pinhas Ben-Tzvi.)

H. Ren and P. Ben-Tzvi are with the Mechanical Engineering Department, Virginia Tech, Blacksburg, VA 24060 USA (e-mail: hailin@vt.edu; bentzvi@vt.edu).

A. Kumar was with the Mechanical Engineering Department, Virginia Tech, Blacksburg, VA 24060 USA. He is now with General Motors Cruise LLC, San Francisco, CA 94103 USA (e-mail: anilks@vt.edu).

Digital Object Identifier 10.1109/JSEN.2019.2893550

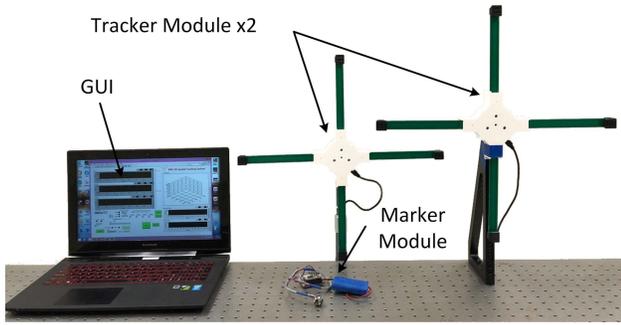


Fig. 1. LOSA-X motion capture system.

accurate enough. To overcome both of these limitations, this paper presents an expandable motion tracking system which offers millimeter level accuracy while being low cost. The proposed tracking system, LOSA-X is built upon previous generation system: the Linear Optical Sensor Array (LOSA) tracking system [21]. Equipped with newer communication architecture, LOSA-X offers higher accuracy, sensing volume, and update rates at a lower cost by simultaneously using multiple trackers. The proposed system uses Extended Kalman Filter (EKF) to fuse multi-sensor data obtained from deployed trackers and active marker. The system also features background illumination removal techniques to estimate and remove background illumination from raw sensor data and isolate the light source of the active marker. This paper presents two trackers-one marker configuration of the proposed system and evaluates its performance.

II. LOSA-X TRACKING SYSTEM IMPLEMENTATION

The LOSA-X motion capture system comprises of multiple tracker modules, a multi-LED illuminated active marker module, and a host computer as shown in Fig. 1. Each module is equipped with a Wi-Fi transceiver module to enable high-speed wireless communication amongst each other. With a 9-DOF IMU and unique line-of-light optics, the proposed system is capable of estimating the orientation and position of the active marker with respect the tracker modules. Similar to LOSA [21], the proposed system makes use of planar stereo triangulation for obtaining the position estimates using the optical sensor arrays as line-scan cameras. The host computer fuses the marker position measurements from each of the tracker modules, along with the inertial measurements from the marker's IMU to track the position and attitude of the marker with high accuracy and precision.

Due to multicast wireless architecture, the proposed system can be operated in either 'computer central' or 'tracker central' modes depending on the final data collection node (computer or tracker). The marker, trackers, and the computer all wirelessly communicate with each other and work in tandem as a single system to estimate the position and attitude of the maker in real-time with high accuracy. In this paper, the computer-centralized architecture has been used for all experiments. In this mode of operation, all trackers and the marker communicate with the host computer directly via a WiFi router for data sharing and operational feedback.

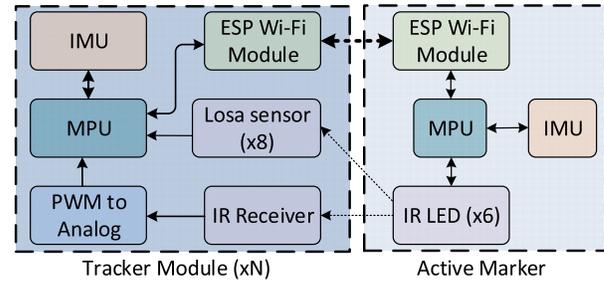


Fig. 2. Hardware schematic of the LOSA tracking system.

Mechanically, the LOSAX tracker modules is 'cross (+)' shaped planer sensing device (made up of a printed circuit board (PCB) substrate) consisting of a central computing circuit with four sensor modules. Each of the four sensor modules are connected to the central computing board via PCB arms, which serve as both mechanical and electrical links. The sensor module enclosure has multiple slots that act as the camera obscura and the camera apertures. Each diagonally opposed sensor module pair forms a stereoscopic line-scan camera pair. The stereoscopic baseline distance is 450 mm for each sensor module pair. The central board has an ABS 3-D printed enclosure, which not only covers and protects it, but also houses a rechargeable battery and helps in mounting the tracker to poles and other rigid structures.

A. Tracker Module Implementation

Each tracker module is controlled using an ARM cortex M4 microcontroller (MCU) mounted on the central computing board. The MCU interfaces with 8 TSL1401R LOSAs (two on each of the four sensor modules), one MPU9250 IMU, and an IR receiver sensor for remote control. The central computing board also hosts an ESP8266 Wi-Fi module (manufactured by Espressif Systems) to enable high-speed communication between the trackers and other systems. Figure 2 shows schematics of the tracker module. Each tracker module is battery-powered (although USB power is possible). The IR (remote) receiver sensor (TSMP6000) is placed at the center of the tracker module to determine the visibility and lighting of the active LED marker. The IR receiver sensor is essentially a band pass filter which responds only to light flashing at high frequencies between 20 and 60 kHz. The receiver responds to flashing light generated by the active marker at this frequency range by returning a pulse width modulation (PWM) output. A multistage passive RC low pass filter is built between the IR sensor and the MCU to obtain the analog signal from the PWM output. The tracker can use this analog signal to detect the presence of a marker.

In order to improve the precision of the position measurement without sacrificing the update rate, the LOSA-X system deploys two TSL1401CL 128-pixel LOSAs in each sensor module. The optical sensor's photodiodes have a peak sensitivity at 780 nm light, which is very close to the infrared LED wavelength used on the active marker.

With the high-speed sampling capability of the MCU's analog to digital convertors (ADC) at 1 Msps, the system takes approximately 600 μ s to read the raw data consisting

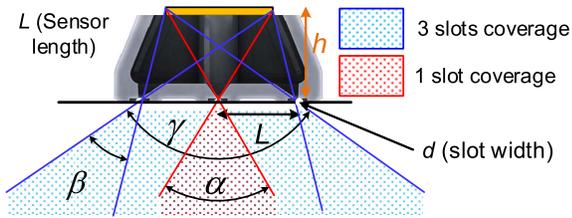


Fig. 3. Sensor module enclosure with multiple slots.

of 1024 data points. The raw data is then passed into software implemented low pass filter (one filter per photodiode) to generate refined data for position estimation. It takes additional $400 \mu\text{s}$ to process the raw data and calculate the position estimate. The sampling period is fixed to 3 ms. The system nodes talk to each other using User Datagram Protocol (UDP) messages, at a communication update rate of nearly 300 Hz. The wireless feedback of the light intensity from the tracker to the marker has also been included to avoid sensor saturation and power inefficiency by adjusting the illumination intensities of the LEDs on the marker.

Similar to the previous iteration [21], a three slot pinhole design was implemented on the LOSA sensor module obscure. As shown in Fig 3, the slot geometry ensures that at any point during operation, only one of the three light beams passing through the cover form an image on the LOSA photosensitive surface. This virtually creates one LOSA sensor for each of the three optical slots and therefore increases the effective sensor resolution threefold. Furthermore, placing two LOSA sensors in opposite directions in one sensor module also increase the robustness of detecting slot changes as the marker moves from one field of view to another. The height of the sensor was set to 12 mm to get field of view, γ , of 90° .

B. Active LED Marker Implementation

The active marker is a small polyhedron, cube in this iteration, with an independently controlled high power, near IR LED on each surface and an MEMS IMU sensor fitted within. The cube structure enables the marker to emit light in any direction using its six LEDs. All of the LEDs are illuminated independently via PWM intensity control mode at a PWM frequency of 30 kHz using a Diode Inc. AL8805 LED driver IC with a >1 MHz switching frequency. By adjusting the PWM duty cycle of each LED, directed illumination can be generated from the active marker at a constant flashing frequency of 30 kHz. The proposed marker uses an InvenSense™ MPU9250® as its IMU to estimate its orientation and position with respect to the tracker modules.

The active marker is connected to a controller board where an ARM Cortex-M4 microcontroller is mounted to regulate the intensity of each LED and process motion data from the IMU. Figure 4 shows the active marker along with its controller board.

III. SYSTEM SELF CALIBRATION AND OPERATION

As the proposed system is designed to work in user defined workspace, there are multiple system parameters which need to be estimated beforehand. Prior to operation the system

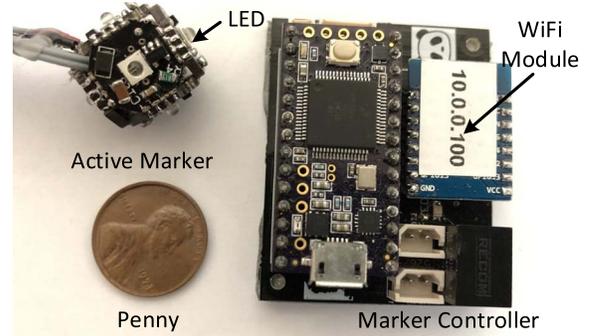


Fig. 4. Active marker with its control board.

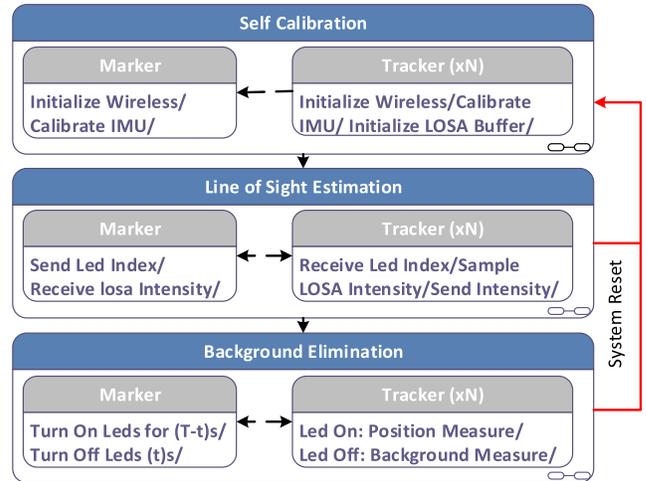


Fig. 5. Operations flowchart of the LOSAX motion tracking system.

performs a self-calibration operation for automatically estimating these parameters. The whole system calibration operation can be divided into three phases: self-calibration, Line of Sight (LoS) estimation, and motion tracking. Figure 5 shows the operational flow chart of the proposed system.

A. System Self Calibration

Once powered on, both the marker and tracker modules enter into the self-calibration phase to automatically estimate system topology and work environment. During this mode, the marker estimates its IMU sensor offsets whereas the trackers estimate both the IMU offsets and the environmental background lighting as seen by optical sensor modules. During this phase, the marker LEDs are turned off and it is assumed that both the trackers and marker remain stationary. Due to this assumption, the mean value obtained for the gyroscope readings from the IMUs correspond to the gyroscopic drift whereas the mean value from the accelerometer represents the gravity vector. These gravity and gyroscope drift vectors are stored for later use during the motion-tracking phase.

Similarly, the scene perceived by the LOSA sensor modules can be assumed to be the background illumination. The line-scan images captured from the LOSA sensor modules are stored in a buffer to estimate the mean background. During this phase, the proposed system also initializes wireless communication between modules and logs their presence into the host PC and the marker by sending UDP messages. Once all

of these processes complete, the tracker modules send UDP messages to the marker to indicate that they are ready for LoS estimation. The self-calibration phase ends when the marker receives a ready message from all trackers.

B. Line of Sight Estimation

Similar to [21], the proposed system uses ‘Line of Sight’ (LoS) vector for directed illumination of the marker and to estimate its attitude with respect to the tracker modules. The LoS vector is defined as a unit vector pointing towards the marker from the tracker center. As the proposed system has multiple tracker modules, LoS vectors for each tracker are estimated and tracked. The LoS is further used for directed illumination of the marker, making it power efficient by powering only the LEDs facing the trackers.

During the LoS estimation phase, the marker turns on one LED on at a time in a predefined order for a fixed duration of time and sends the LED status to the tracker modules. The tracker modules then send back the received LED light intensity as a feedback. Based on this LED intensity feedback, the marker computes the LoS vector for each of the tracker modules. Assuming N tracker modules in the LOSA system, the LoS vector to any j th tracker module can be determined on the basis of LED index, i , and intensity feedback, I_{ij} , received from the concerned module. Once the maximum intensities from all trackers for the six LEDs on the marker are received, the LoS vector can be estimated. The LoS vector (\vec{r}_j) for the j th module can be calculated as,

$$\vec{r}_j = \left(\sum_i I_{ij} \vec{V}_i \right) / \left(\sum_i I_{ij} \right) \quad (1)$$

where, \vec{V}_i represents the outward illumination vector for the i th LED on the marker.

C. Background Elimination

As the LOSA sensor modules are sensitive to a broad spectrum of light, the ambient/background illumination may severely affect detection of the active marker in the line scan images. The proposed sensor system makes novel use of IR remote receivers to identify and subtract background illumination from the LOSA images. The background illumination removal not only eliminates the need for IR-pass filters (as used on the previous generation) for identification of markers in the images, but also enables the LOSA-X system to perform in outdoor environments with multiple or complex sources of illumination. This process is completed through cooperation between the trackers and markers.

The background elimination process consists of two cyclically repeating operational modes: background sensing and background removal. The background-sensing mode lasts only for a fraction of the time of the position-sensing mode. During this mode, the active marker turns off the LED illumination, which is detected by the tracker through the IR receiver at the onset of the background-sensing mode. In this mode, the tracker starts filling and updating an adjustable depth LOSA image buffer with five images per LOSA to capture the background illumination as perceived by each LOSA pixel.

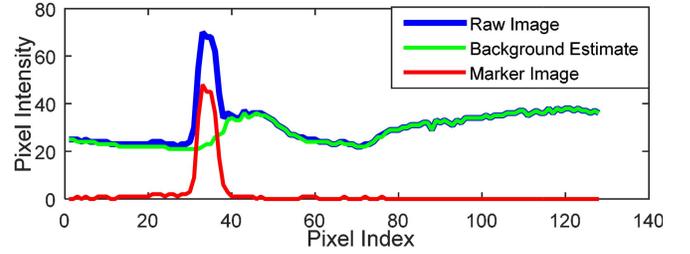


Fig. 6. Background subtraction in LOSA image.

The background illumination image is computed taking pixel-wise mean of LOSA background buffers. As the system transitions from the background sensing to the background removal mode, the foreground-marker image is computed by pixel-wise subtraction of the most recent background illumination image from the most recent acquired LOSA image. As the trackers are assumed to be stationary during operation, the static background doesn't change and can be estimated and corrected for reliably. Figure 6 shows the extracted foreground marker image for one of the LOSA sensors during the background illumination process.

D. Position and Attitude Measurement

Each tracker module in the LOSA-X system can estimate the marker position in its frame of reference by using planer stereo triangulation [21]. For each LOSA module pair, the 2D position of the marker (X, Z) can be estimated as follows:

$$X = b \frac{x_A + x_B}{2(x_A - x_B)}, \quad Z = b \frac{h}{x_A - x_B} \quad (2)$$

where b represents the baseline distance between the stereoscopic LOSA pairs and h is the height of the aperture slot above the LOSA sensor surface. Also, x_A and x_B represent the position of the marker's light peak in background eliminated marker images from the stereoscopic sensor pair. The sub-pixel accurate marker peak position was estimated taking weighted mean of the image indices in a 7-pixel wide window (empirically optimized) with weights being the intensity value in the background removed marker image. With four LOSAs in each stereovision pair (two on each side), there are four stereo-pair combinations possible in each sensing plane. Position estimates were computed for all the four stereo-pair combinations and the average value of the 2D position coordinates was retained and used for motion tracking.

By fusing 2D position estimates from the horizontal and vertical stereoscopic pairs, the 3D position of the marker can be computed. Along with estimating the position of the marker, each tracker continuously updates the attitude of the marker in its own frame using LoS and the gravity vectors. Note that the LoS vector of each tracker module coincides with the initial position estimate vector from each module. The rotation matrix from any tracker T_i to marker m can be obtained as follows,

$$\mathbf{T}_{T_i}^m = \begin{bmatrix} \hat{\mathbf{r}}_i & \hat{\mathbf{m}}\mathbf{g} & (\hat{\mathbf{r}}_i \times \hat{\mathbf{m}}\mathbf{g}) \end{bmatrix} \begin{bmatrix} \hat{\mathbf{T}}_i\mathbf{p} & \hat{\mathbf{T}}_i\mathbf{g} & (\hat{\mathbf{T}}_i\mathbf{p} \times \hat{\mathbf{T}}_i\mathbf{g}) \end{bmatrix}^{-1} \quad (3)$$

where, \mathbf{r} represents LoS, \mathbf{g} represent gravity vector, \mathbf{p} represents the 3D position vector of the marker, and $\hat{\cdot}$ represents

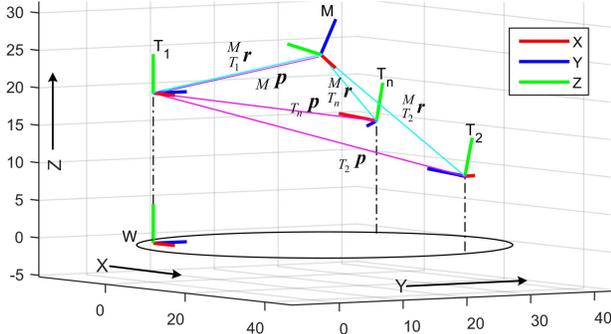


Fig. 7. Coordinate system for generic multi-tracker LOSA-X topology.

the unit vector normalization operation. For computational convenience, the attitude estimates are stored and computed in quaternion form whenever possible. The rotation matrix T can also be written in terms of the unit quaternion vector $\mathbf{q} = [q_0 \ q_1 \ q_2 \ q_3]^T$ as

$$\begin{aligned} \mathbf{T}_s^m(\mathbf{q}) &= \begin{bmatrix} 1 - 2(q_2^2 + q_3^2) & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & 1 - 2(q_1^2 + q_3^2) & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & 1 - 2(q_1^2 + q_2) \end{bmatrix} \end{aligned} \quad (4)$$

The quaternion, \mathbf{q} , is updated by integrating the angular rates on the microcontroller that are then sent to the PC for data fusion.

IV. MULTI TRACKER FUSION

In the proposed multiple tracker system, each tracker can provide a pose estimation of the active marker in the form of Cartesian coordinates and quaternions. The trackers triangulate the marker using the LOSA sensors while the orientation information of the marker is obtained from the IMU attached to the marker. The LOSA-X tracking system fuses inertial data from the marker with the position estimates from each of the tracker modules for robust motion tracking. An Extended Kalman Filter (EKF) has been applied to solve the nonlinear state system to provide more robust and stable pose estimation of the active marker.

A. System Modeling

At the lowest level, each tracker module in the system is designed to fuse optical position estimates from LOSA sensors and inertial data from the active marker to obtain the pose of the marker in the local frame of reference. Out of all the tracker modules in the system, one tracker is assigned as the master node by assigning its coordinate frame of reference as a global frame of reference. At a higher system level, pose estimates from each tracker are fused to obtain the position and pose of the marker in the master tracker module frame with better accuracy than a single tracker. Figure 7 shows the generic coordinate system for the proposed system, where cyan and magenta colored rays represent LoS vectors (\mathbf{r}) and 3D position vectors (\mathbf{p}) respectively for each of the tracker modules. Similar to the implementation in [21], the EKF is deployed on the PC in LabVIEW® while raw data collection,

low pass filters and high speed wireless communication are running on the microprocessors and WiFi modules. Using a similar modeling method that was applied in [32], the nonlinear stochastic processes of the LOSA-X system can be expressed as,

$$\begin{aligned} \mathbf{x}_k &= f(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{w}_{k-1} \\ \mathbf{y}_k &= h(\mathbf{x}_k) + \mathbf{v}_k \end{aligned} \left\{ \begin{array}{l} \mathbf{w}_k \in N(\vec{0}, \mathbf{Q}_k), \\ \mathbf{v}_k \in N(\vec{0}, \mathbf{R}_k) \end{array} \right. \quad (5)$$

where, f , h are the state transition and measurement models, \mathbf{x}_k is the state vector of the system at time step k , \mathbf{u}_k is the control input vector applied to the system, \mathbf{w}_k and \mathbf{v}_k are the process and measurement noise vectors of the nonlinear system, respectively. The noise vectors \mathbf{w}_k and \mathbf{v}_k are assumed to follow multivariate Gaussian distributions where \mathbf{Q}_k and \mathbf{R}_k are the covariance matrices of those distributions, respectively.

In the propose system model, the state vector, \mathbf{x}_k , is represented in terms of the position, \mathbf{p} , velocity, \mathbf{v} , and unit rotation quaternion, \mathbf{q} , of the marker with respect to each of the tracker modules as follows:

$$\mathbf{x}_k = [\mathbf{T}_1\mathbf{p}_k \quad \mathbf{T}_1\mathbf{v}_k \quad \mathbf{T}_1\mathbf{q}_k \quad \mathbf{T}_i\mathbf{p}_k \quad \mathbf{T}_i\mathbf{q}_k]_{2 \leq i \leq N} \quad (6)$$

where, N and T_1 represent the number of tracker modules in the system topology and master tracker module, respectively. The linear acceleration of the active marker is chosen as the input vector to the system. To express the acceleration in the main tracker frame, the rotation matrix between the marker and main tracker is updated through the measured angular velocity from the IMU on the marker. The input of the system at time step k , \mathbf{u}_k , can be expressed as,

$$\mathbf{u}_k = \mathbf{T}(\mathbf{T}_1\mathbf{q}_{k-1})\mathbf{M}\mathbf{a}_k - \mathbf{T}_1\mathbf{a}_k \quad (7)$$

where $\mathbf{T}(\mathbf{q})$ represents the rotation matrix (based on most recent attitude estimates) mapping quantities in the marker's frame of reference to the master tracker's frame and \mathbf{a} represents the measured acceleration vector from the IMU on the marker, M , and the master tracker, T_1 , module, respectively. The state transition model can be described as follows:

$$\begin{bmatrix} \mathbf{M}\mathbf{p}_k \\ \mathbf{M}\mathbf{v}_k \\ \mathbf{M}\mathbf{q}_k \\ \mathbf{T}_i\mathbf{p}_k \\ \mathbf{T}_i\mathbf{q}_k \end{bmatrix} = \begin{bmatrix} \mathbf{M}\mathbf{p}_{k-1} + \mathbf{M}\mathbf{v}_{k-1} \cdot \delta t + 0.5\mathbf{u}_k \delta t^2 \\ \mathbf{M}\mathbf{v}_{k-1} + \mathbf{u}_k \delta t \\ \mathbf{M}\mathbf{q}_{k-1} + \dot{\mathbf{q}}(\mathbf{M}\mathbf{q}_k, \mathbf{M}\boldsymbol{\omega}_{k-1}) \cdot \delta t \\ \mathbf{T}_i\mathbf{p}_{k-1} \\ \mathbf{T}_i\mathbf{q}_{k-1} \end{bmatrix}_{1 < i \leq N} \quad (8)$$

where δt is the interval between two consecutive time steps, and the quaternion rate, $\dot{\mathbf{q}}(\mathbf{q}, \boldsymbol{\omega})$, is updated from the measured angular rates $\boldsymbol{\omega} = [\omega_x \ \omega_y \ \omega_z]$ by

$$\dot{\mathbf{q}}(\mathbf{q}, \boldsymbol{\omega}) = \frac{1}{2} \begin{bmatrix} q_0 & q_3 & -q_2 & q_1 \\ q_1 & q_2 & q_3 & -q_0 \\ q_2 & -q_1 & q_0 & q_3 \\ q_3 & -q_0 & -q_1 & -q_2 \end{bmatrix} \begin{bmatrix} 0 \\ \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \quad (9)$$

The measured position, velocity and acceleration of the marker are chosen as the observable system variables.

The observation model of the system can be written as follows:

$$\begin{bmatrix} M\mathbf{r}_k \\ M\mathbf{v}_k \\ M\mathbf{a}_k \\ M_i\mathbf{r}_k \\ M_i\mathbf{v}_k \\ M_i\mathbf{a}_k \end{bmatrix} = \begin{bmatrix} M\mathbf{p}_k \\ M\mathbf{v}_k \\ \frac{M\mathbf{v}_k - M\mathbf{v}_{k-1}}{\delta t} + \mathbf{g} \\ \mathbf{T}(T_i\mathbf{q}_k)^T (M\mathbf{p}_k - T_i\mathbf{p}_k) \\ \mathbf{T}(T_i\mathbf{q}_k)^T M\mathbf{v}_k \end{bmatrix} \quad 1 < i \leq N \quad (10)$$

B. EKF Implementation

The sensor fusion in the proposed system was implemented using Extended Kalman Filters (EKF) on the basis of the state transition and observation model shown in equation (8) and (10). In each iteration, the state estimates and covariance estimate are updated using prediction and correction (innovation) stages of the EKF as follows:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}, & \mathbf{P}_{k|k} &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \\ \hat{\mathbf{x}}_{k|k-1} &= f(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_{k-1}), & \mathbf{P}_{k|k-1} &= \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_k \end{aligned} \quad (11)$$

where,

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{y}_k - h(\hat{\mathbf{x}}_{k|k-1}), & \{\mathbf{F}_k, \mathbf{H}_k\} &= \left\{ \frac{\partial f}{\partial \mathbf{x}}, \frac{\partial h}{\partial \mathbf{x}} \right\} \Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}}, \\ \mathbf{K}_k &= \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1} & \text{and } \mathbf{S}_k &= \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k. \end{aligned} \quad (12)$$

Here, the variables $\tilde{\mathbf{y}}$, \mathbf{S}_k , \mathbf{F}_k , \mathbf{H}_k , \mathbf{K}_k , and \mathbf{I} represent the innovation residuals, innovation covariance matrix, state transition noise covariance matrix, observation noise covariance matrix, Kalman gain matrix, and identity matrix, respectively.

V. EXPERIMENTS AND EVALUATION

A. Background Light Immunity

To test the background rejection capability and adaptability to different lighting conditions of the proposed system, portable light sources are used to imitate dynamically varying bright light environments. Different lighting patterns are applied to prove the system's adaptability in both diverse static lighting conditions and dynamic lighting conditions. Figure 8 shows three samples of marker images captured while changing the background illumination. The peak value after the background removal process in all situation ranges from 9 to 12 pixels which is bright enough to be detected reliably by the system.

B. Motion Tracking Capability

To evaluate the position estimation capability of the proposed system, the motion of the marker was tracked using a two-tracker (namely *Tracker1* and *Tracker2*) LOSA-X system prototype. In the experimental setup, the marker is mounted on a rod hinged on an optical table and constrained to move in plane on a circular trajectory of radius 260 mm with respect to the hinge position. The tracker modules are placed at different locations with respect to the marker module with the position accurately measured using laser range finders. As the motion of the marker is constrained to a circle and with relative position known, the estimated position estimates of the marker can be

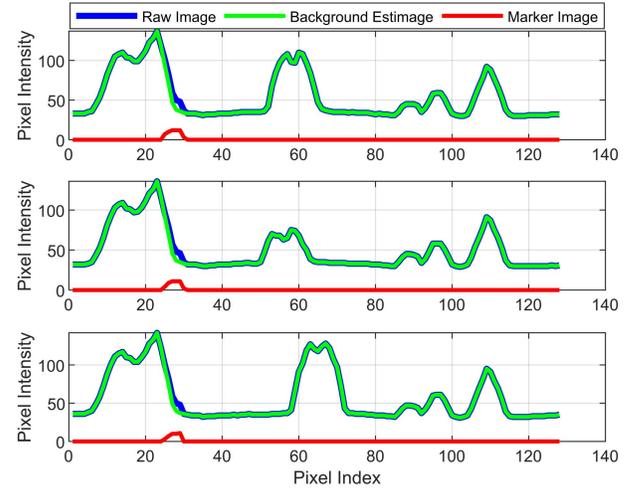


Fig. 8. Dynamic bright background illumination removal.

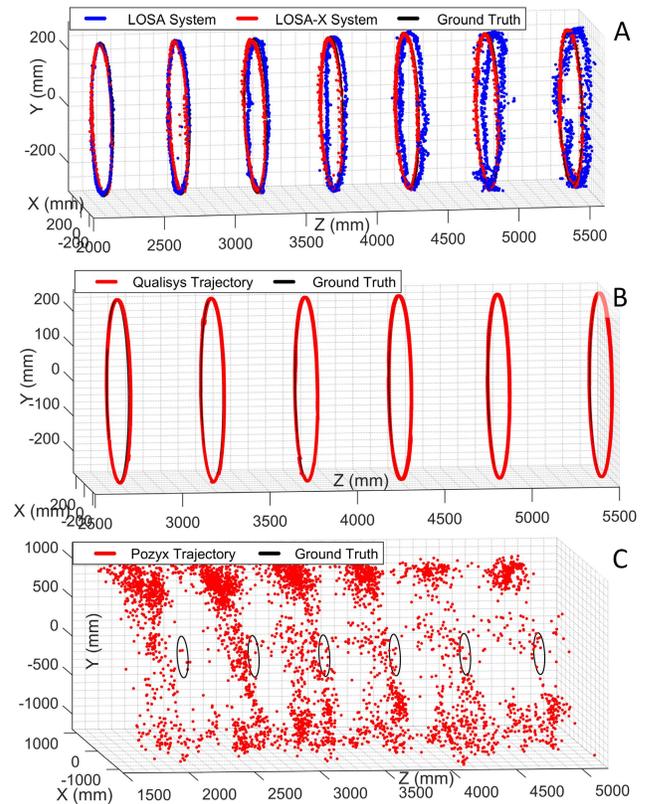


Fig. 9. Position measurement with Pozyx and Qualisys.

compared against the expected trajectory for system evaluation while rotating the marker in a circular motion. Figure 9A compares the position estimates from system with the expected trajectory at different operating ranges. The black circles show the ground truth; blue and red dots show the estimates from *LOSA System* (single tracker) [21] and the proposed *LOSA-X* (2 tracker configuration), respectively; the blue dots show the fused position estimates. Figures 9B and 9C show similar trajectory results obtained from commercially available motion tracking systems Qualisys [26] (in 10-camera configuration) and Pozyx [17], respectively.

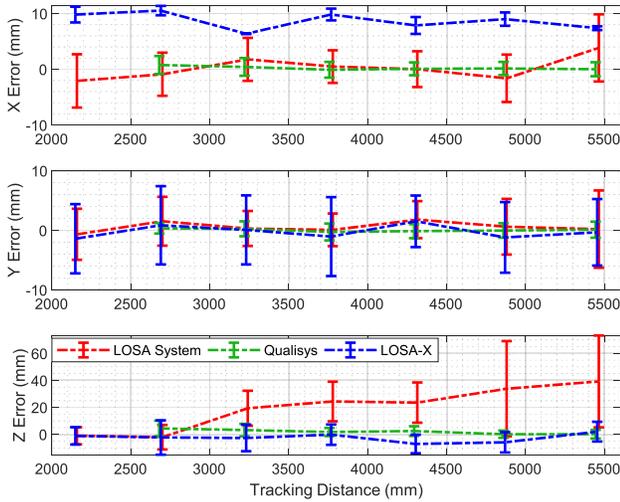


Fig. 10. Position accuracy of proposed LOSA-X system, LOSA and Qualisys tracking systems at different operating ranges.

TABLE I

SENSORS SPECIFICATIONS AND PERFORMANCE EVALUATION

System Parameter	LOSA System	Pozyx Tracker	Qualisys Tracker	LOSA-X System
Accuracy (mean error)	6.8 mm*	47.2 mm*	0.8 mm*	1.9 mm*
Repeatability (error std. dev.)	8.8 mm*	415.9mm*	2.1 mm*	4.6 mm*
Min. Tracker units required	1	4+1	4 [†]	2
Maximum Range	6 m*	30-10 m [†]	35m [†]	>6m, (Expandable)
Max Frame Rate	130 Hz*	140 Hz [†]	1752 Hz	<300 Hz
Approx. Cost (USD)	~70	680	>10,000	~130

*Experimental Data. [†]From studies/datasheets.

C. Performance Evaluation

For quantitative analysis, mean absolute value and standard deviation of the position error were observed in experiments conducted at different operating distances. Figure 10 shows the mean error and standard deviation of position estimates from LOSA System, Qualisys tracking system and the proposed LOSA-X System. Compared to single tracker measurements (LOSA system), fused multi-tracker LOSA-X position estimates provide better tracking accuracy. Up to 55% reduction in position estimation error was achieved using the proposed sensor fusion technique over the previous generation LOSA tracking system. Commercial products such as Pozyx [17] and Qualisys [26] were also evaluated in similar experiments for performance comparison. Qualisys system (equipped with 10 cameras) delivered excellent motion tracking accuracy with high repeatability. Motion tracking accuracy from the Pozyx motion tracker was poorer than expected and hence was not plotted in Figure 10. We believe that the reason behind the poor performance of Pozyx is the multipath reflections (in indoor environment) along with the use of directional antennas of the sensor. Table 1 shows system specifications of the proposed LOSA-X and other sensors used in experiments.

To assess the accuracy of the attitude estimates obtained from the LOSA-X system, a commercially available

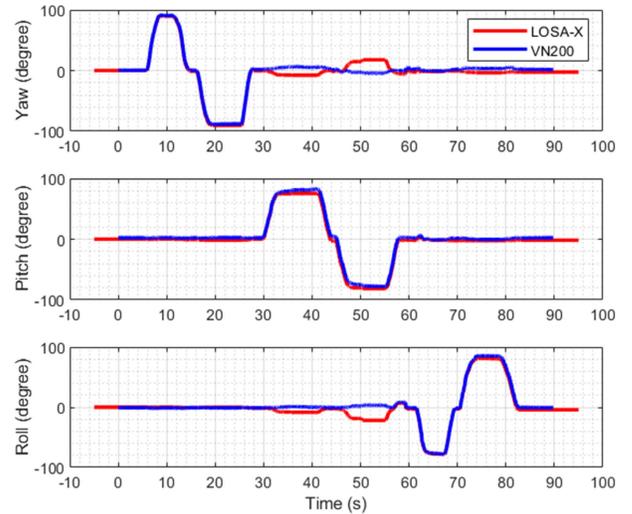


Fig. 11. Orientation measurement with VN200 INS and the proposed LOSA-X tracking system.

VN200 INS/AHRS sensor is attached to the active marker as a reference. In this experiment, the marker, along with the VN200, is rotated around each axis by approximately 90° in a clear space to obtain orientation data. Figure 11 shows the comparison of roll, pitch and yaw estimates from both the VN200 and LOSA-X system. Attitude estimates from the proposed sensor closely matched with those from the VN200 sensor.

VI. CONCLUSIONS AND FUTURE WORK

In this work, we presented a novel multi-sensor expandable motion tracking system with significant upgrades over its previous generation tracking system. The proposed tracking system is able to reliably reject dynamic background illumination, thereby eliminating the need for an IR filter and enabled the system to operate outdoor. Furthermore, the use of multiple tracker modules enabled the system to expand the tracking volume as per operational needs. This also increases the accuracy of the system up to 55% compared to the previous generation system. Furthermore, the motion tracking accuracy of the proposed system was estimated to be comparable to a 100 times more expansive state of the art motion tracking system. With the incorporation of Wi-Fi communication, this system enables high speed communication between each module. All of these changes allow for the proposed system to achieve high precision low-cost 3D tracking for a wide range of indoor and outdoor applications. Multiple experiments validated performance improvements introduced in this system.

Currently, position calibration is performed for each tracking module to refine the position estimation; however, this is not suitable for mass production. Automated assembly and better machining techniques should mitigate some of these limitations. Currently, the tracking module and marker module shift most of the computational burden to the PC due to the limits on the computing ability of the currently deployed microprocessor. This increases communication delay within the system. A more powerful microprocessor/embedded

computer would remove the dependence on the PC for computation and allow for simultaneous data fusion with a large number of trackers in the system. In both the LOSA/LOSA-X tracking systems, only one active marker is allowed to be tracked at a time. This is due to the single spectrum of light used in the tracking system where no identification pattern is deployed to distinguish between different markers. However, tracking independent objects at the same time is a requirement in many applications. Time multiplexing can be introduced into the system for tracking multiple markers at the cost of lower update frequency.

Ongoing and future work to upgrade the LOSA-X tracking system involves dedicated manufacturing methods to create the mechanical parts of LOSA-X. Circuit board construction and assembly may be performed on SMT assembly station using IPC-A-610 Class 2 standard. ARM Cortex M7 microprocessors will be used in future versions to allow for faster computation. A 16-Channel, 16-Bit ADC, such as the ADS1162, could be used to sample LOSA sensors at 23.7kSPS/channel for faster data acquisition and higher dynamic range.

ACKNOWLEDGMENT

The authors would like to acknowledge the help from Dr. Robin Queen of Virginia Tech in facilitating motion tracking experiments with the Qualisys System.

REFERENCES

- [1] M. Buck *et al.*, "Controlled flight of a biologically," *Science*, vol. 80, pp. 603–607, May 2013.
- [2] A. Kushleyev, D. Mellinger, C. Powers, and V. Kumar, "Towards a swarm of agile micro quadrotors," *Auton. Robots*, vol. 35, no. 4, pp. 287–300, 2013.
- [3] V. Kumar and E. Todorov, "MuJoCo HAPTIX: A virtual reality system for hand manipulation," in *Proc. IEEE-RAS Int. Conf. Hum. Robot.*, Nov. 2015, pp. 657–663.
- [4] M. C. Johnson-Glenberg, D. A. Birchfield, L. Tolentino, and T. Koziupa, "Collaborative embodied learning in mixed reality motion-capture environments: Two science studies," *J. Educ. Psychol.*, vol. 106, no. 1, pp. 86–104, 2014.
- [5] J. C. P. Chan, H. Leung, J. K. T. Tang, and T. Komura, "A virtual reality dance training system using motion capture technology," *IEEE Trans. Learn. Technol.*, vol. 4, no. 2, pp. 187–195, Apr./Jun. 2011.
- [6] L. Vera, J. Gimeno, I. Coma, and M. Fernández, "Augmented mirror: Interactive augmented reality system based on kinect," in *Human-Computer Interaction—INTERACT*. Berlin, Germany: Springer, 2011, pp. 483–486.
- [7] A. Fernández-Baena, A. Susfín, and X. Lligadas, "Biomechanical validation of upper-body and lower-body joint movements of Kinect motion capture data for rehabilitation treatments," in *Proc. 4th Int. Conf. Intell. Netw. Collaborative Syst. (INCoS)*, Sep. 2012, pp. 656–661.
- [8] L. Sigal, A. O. Balan, and M. J. Black, "HumanEva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion," *Int. J. Comput. Vis.*, vol. 87, nos. 1–2, pp. 4–27, 2010.
- [9] V. B. Zordan, A. Majkowska, B. Chiu, and M. Fast, "Dynamic response for motion capture animation," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 697–701, 2005.
- [10] E. Sifakis, I. Neverov, and R. Fedkiw, "Automatic determination of facial muscle activations from sparse motion capture marker data," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 417–425 2005.
- [11] K. Aminian and B. Najafi, "Capturing human motion using body-fixed sensors: Outdoor measurement and clinical applications," *Comput. Animation Virtual Worlds*, vol. 15, no. 2, pp. 79–94, 2004.
- [12] E. Kelasidi, P. Liljeback, K. Y. Pettersen, and J. T. Gravdahl, "Innovation in underwater robots: Biologically inspired swimming snake robots," *IEEE Robot. Autom. Mag.*, vol. 23, no. 1, pp. 44–62, Mar. 2016.
- [13] A. Schmitz, M. Ye, G. Boggess, R. Shapiro, R. Yang, and B. Noehren, "The measurement of *in vivo* joint angles during a squat using a single camera markerless motion capture system as compared to a marker based system," *Gait Posture*, vol. 41, no. 2, pp. 694–698, 2015.
- [14] E. L. Secco, A. T. Maereg, D. Reid, and A. K. Nagar, "An integrated haptic system combining VR, a markerless motion capture system and tactile actuators," *ICST Trans. Ambient Syst.*, vol. 5, no. 17, p. 154375, Mar. 2018, doi: [10.4108/eai.23-3-2018.154375](https://doi.org/10.4108/eai.23-3-2018.154375).
- [15] L. Xu *et al.*, "FlyCap: Markerless motion capture using multiple autonomous flying cameras," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 8, pp. 2284–2297, Aug. 2018.
- [16] Y. Wang, Y. Liu, X. Tong, Q. Dai, and P. Tan, "Outdoor markerless motion capture with sparse handheld video cameras," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 5, pp. 1856–1866, May 2018.
- [17] Pozyx Labs, Ghent, Belgium. *Pozyx—Accurate Positioning*. Accessed: Jan. 24, 2019. [Online]. Available: <https://www.pozyx.io/>
- [18] T. Cloete and C. Scheffer, "Benchmarking of a full-body inertial motion capture system for clinical gait analysis," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 4579–4582.
- [19] A. G. Kirk, J. F. O'Brien, and D. A. Forsyth, "Skeletal parameter estimation from optical motion capture data," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2005, pp. 782–788.
- [20] A. Pfister, A. M. West, S. Bronner, and J. A. Noah, "Comparative abilities of microsoft Kinect and vicon 3D motion capture for gait analysis," *J. Med. Eng. Technol.*, vol. 38, no. 5, pp. 274–280, 2014.
- [21] A. Kumar and P. Ben-Tzvi, "Spatial object tracking system based on linear optical sensor arrays," *IEEE Sensors J.*, vol. 16, no. 22, pp. 7933–7940, Nov. 2016.
- [22] T. Monnet, M. Samson, A. Bernard, L. David, and P. Lacouture, "Measurement of three-dimensional hand kinematics during swimming with a motion capture system: A feasibility study," *Sport. Eng.*, vol. 17, no. 3, pp. 171–181, 2014.
- [23] R. Raskar *et al.*, "Prakash: Lighting aware motion capture using photosensing markers and multiplexed illuminators," *ACM Trans. Graph.*, vol. 26, no. 3, Jul. 2007, [Online]. Available: <http://doi.acm.org/10.1145/1276377.1276422>, doi: [10.1145/1276377.1276422](https://doi.org/10.1145/1276377.1276422).
- [24] NaturalPoint, Inc., Corvallis, OR, USA. *Optitrack*. Accessed: Jan. 24, 2019. [Online]. Available: <https://optitrack.com/>
- [25] Vicon Motion Systems Ltd., Oxford, U.K. *Motion Capture Systems*. Accessed: Jan. 24, 2019. [Online]. Available: <https://www.vicon.com/>
- [26] Qualisys AB, Kvarnbergsgatan, Göteborg, Sweden. *Qualisys*. Accessed: Jan. 24, 2019. [Online]. Available: <https://www.qualisys.com/>
- [27] A. Aristidou and J. Lasenby, "Motion capture with constrained inverse kinematics for real-time hand tracking," in *Proc. 4th Int. Symp. Commun., Control Signal Process. (ISCCSP)*, Mar. 2010, pp. 1–5.
- [28] K. Zhou, X. Wang, H. Wei, L. Yin, Z. Wan, and Z. Wang, "Modeling and calibration of a precise optical positioning system based on four linear cameras," *Appl. Opt.*, vol. 57, no. 19, pp. 5538–5548, 2018.
- [29] H. Liu, L. Yang, Y. Guo, R. Guan, and J. Zhu, "Precise calibration of linear camera equipped with cylindrical lenses using a radial basis function-based mapping technique," *Opt. Express*, vol. 23, no. 3, pp. 3412–3426, Feb. 2015.
- [30] Q. Wen and J. Wu, "Linear CCD based optical tracking using stereo correspondence algorithm," in *Proc. Int. Conf. Artif. Intell. Comput. Intell.*, Nov. 2009, pp. 422–425.
- [31] Z. Zhang, "Microsoft Kinect sensor and its effect," *IEEE MultiMedia*, vol. 19, no. 2, pp. 4–10, Feb. 2012.
- [32] M. B. Alataise and G. P. Hancke, "Pose estimation of a mobile robot based on fusion of IMU data and vision data using an extended Kalman filter," *Sensors*, vol. 17, no. 10, p. 2164, 2017.

Hailin Ren, photograph and biography not available at the time of publication.

Anil Kumar, photograph and biography not available at the time of publication.

Pinhas Ben-Tzvi, photograph and biography not available at the time of publication.