Spatial Object Tracking System Based on Linear Optical Sensor Arrays

Anil Kumar, *Student Member, IEEE*, and Pinhas Ben-Tzvi, *Senior Member, IEEE*

***Abstract*— This paper presents a low-cost novel motion tracking system using linear optical sensor arrays supplemented by a 9DOF inertial measurement unit. Each pair of the linear sensor modules is capable of tracking the 2D location of an active infrared LED light source (marker). The presented tracking system combines two of these modules to realize spatial position tracking of the LED light source. The proposed system efficiently amalgamates inertial data with optical tracking estimates to extract attitude, minimize power requirement, and improve positional accuracy. The tracking system enables high-speed, high-accuracy, and low-cost position and attitude tracking for a wide range of applications. The detailed hardware design and performance evaluation experiments of this approach are also reported. The performance of the proposed tracking system was evaluated by measuring the tracking error with respect to known trajectories of the LED light source from different distances.**

***Index Terms*— 3D tracking, linear optical sensor array, optical tracking, extended Kalman filter (EKF), inertial measurement unit (imu), planar stereo vision.**

I. Introduction

THREE-DIMENSTIONAL tracking plays a vital role in various engineering disciplines involving motion analysis in industry and academia, and a significant area of research is developing low-cost, accurate tracking sensors. Certain engineering applications like robotics, navigation, biomechanics and virtual reality often require real-time position tracking [1]–[5]. Depending on the application, the position tracking systems are selected based on measurement range, accuracy and update rate. Position tracking systems can be broadly categorized into two types: contact-based and contact-less. This paper will limit the discussion to contact-less tracking systems.

Contact-less tracking systems often measure electromagnetic/acoustic modalities without making physical contact for tracking purposes. Based on the type of the modality used, major tracking systems include Magnetic Tracking [6]–[8], Inertial Tracking [9], [10], RSSI Tracking [11], [12] and Optical Tracking. The scope of the paper is limited to optical tracking.

Optical trackers work through triangulation and require direct line-of-sight from the camera to the measurement point. They are also capable of larger work envelopes. Optical tracking methods are categorized into two groups: marker-less and marker-based tracking systems.

A. Marker-Less Camera-Based Tracking

The simplest marker-less optical tracking is a calibrated monocular camera tracking. If the camera is rigidly mounted normal to the plane of motion, then the pixel coordinates of an object being tracked can be mapped to the actual 2D position in plane. Spatial optical tracking often requires multiple perspectives (like stereovision) of the tracked object. Stereovision requires simultaneous identification and tracking of an object in two (or more) images, which is computationally expensive. Spatial object positioning is obtained from the disparity in position of the object in planar images [13], [14]. Marker-less visual tracking systems have the potential to provide natural, non-contact methods to measure object position and orientation. Although powerful and accurate, a camera-based tracking system is computationally expensive, leading to a relatively low update rate.

B. Marker-Based Tracking

Marker-based tracking systems [15]–[17] typically consist of surface markers and image sensors (e.g., the Vicon motion systems). Specialized optical filters are often used to isolate markers from the background and significantly reduce computational load. These systems offer higher accuracy and higher update rates compared with the marker-less camera-based tracking systems. Such commercial tracking systems typically cost thousands of dollars and require time-consuming calibration.

Time multiplexed tracking systems like Microsoft Kinect, MEMSEye and Lidar project their own marker patterns and use optical triangulation to obtain 3D location [18], [19]. These sensors are less expensive and can achieve moderate to high update rates but at the expense of accuracy. Low cost tracking systems such as Lumitrack [20] project static laser patterns and use the patterns to track relative motion. These sensors are adept at measuring relative velocity between the base frame and moving object with high update rates, but don't provide accurate position estimates. In a similar approach, Bokode [21] projects position information in the form of QR codes embedded on optical bokeh. The position information of 'Bokodes' is then used to estimate the position and orientation of the (SLR) camera receiver with respect to the bokodes.

In contrast to these tracking systems, the proposed linear optical sensor array (LOSA) tracking system offers high

Manuscript received May 9, 2016; revised August 31, 2016; accepted August 31, 2016. Date of publication September 8, 2016; date of current version October 13, 2016. The associate editor coordinating the review of this paper and approving it for publication was Dr. M. N. Abedin.

A. Kumar is with the Mechanical Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061 USA (e-mail: anilks@vt.edu).

P. Ben-Tzvi is with the Mechanical Engineering Department and the Electrical and Computer Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061 USA (e-mail: bentzvi@vt.edu).

Digital Object Identifier 10.1109/JSEN.2016.2607120

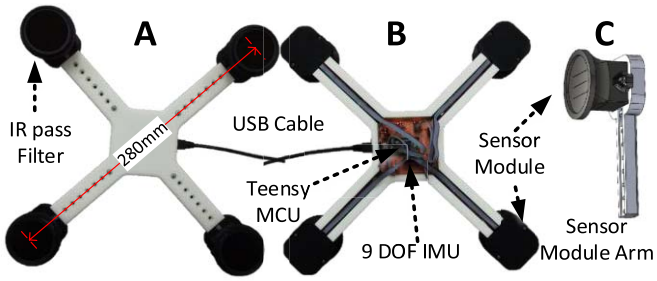

Fig. 1. LOSA 3D tracking system. (A) Top View. (B) Bottom View with cover removed. (C) CAD of the sensor module arm.

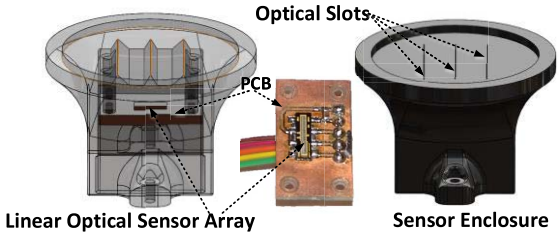

Fig. 2. CAD model of the LOSA sensor assembly.

accuracy and update rates at low cost. The proposed LOSA tracking system tracks an active marker (omnidirectional IR LED tetrahedron) at millimeter-level accuracy with very high update rate (∼250 fps). The proposed system combines inertial and optical tracking without relying on magnetometer data, which is susceptible to stray magnetic noise and provides more reliable position and attitude estimates. Experimental results demonstrate millimeter-level accuracy in the tracking system's operational range. Furthermore, the system's total cost including components and fabrication is less than $50 when manufactured in large volume. Finally, the system is relatively small, lightweight and low powered, making it suitable for wearable and/or mobile devices. The presented sensor when combined with other input/output devices offers significant potential for a range of applications such as mobile robot tracking, human computer interaction and virtual reality.

Use of line cameras in computer vision is not new. Researchers have been using linear cameras for computer vision for quite some time [22]–[25]. This paper presents an application of a low cost LOSA in coordination with an IMU to obtain position and attitude estimates.

II. Tracking System Implementation

The key components of the presented tracking system are LOSA sensor modules and Active LED maker equipped with a 9-axis IMU and wireless connectivity. The tracking system locates the LED marker by using triangulation on a pinhole camera image of a marker on a photodiode array area (length). The system estimates triangulation disparity by finding the peak value positions of the signals in the sensor array while adjusting for diffraction effects. Thus, each pair of sensor modules facilitates 2D position tracking in the epipolar plane of the sensor pair. Utilizing multiple sensor pairs, the marker's spatial position may be extrapolated from multiple planar measurements. A computer then fuses the position estimates from stereovision and IMU data using EKF to extract more accurate position and attitude estimates.

A. LOSA Sensor Mechanical Details

The core of the proposed tracing system is the LOSA sensor modules with optical slots. The sensor comprises of 4 LOSA sensor modules each connected to the central chassis (white) through adjustable arms (Fig. 1). The sensor module enclosure acts a camera obscura and the optical slots act as camera apertures (Fig. 2). Each sensor module enclosure fits onto a 'sensor module arm'. The precise position of the sensors with respect to each other is ensured through the 3D printed (ABS plastic) chassis. In fully retracted condition, the stereoscopic baseline distance is 280mm for each sensor module pair. Furthermore, the guiding rail structure allows arm extension at 10 mm intervals, thus making the baseline adjustable to a maximum of 400mm for long range operations. To filter out noise due to ambient visible light, an IR-pass filter (850nm) was mounted on top of each sensor module.

B. Sensor Implentation

Each LOSA sensor module consists of a custom printed circuit board, which carries the linear optical sensor array, a microcontroller, a cover box and visible light filter. Each module has a TSL1401R 128-pixel linear array sensor manufactured by AMS AG. When two sensor modules are mounted and aligned with each other (Fig. 1), this configuration enables planar sensing through triangulation. Compared to a 2D camera sensor, this setup greatly reduces total pixels to process (128×2 pixels vs 128×128 pixels) and thus dramatically increases the update rate. The optical sensor's photodiodes have a highly sensitive response to a broad spectrum of light and the peak responsivity wavelength of the photodiode is at around 800 nm, which is very close to the infrared LED wavelength used. An ARM Cortex-M4 based 96 MHz Teensy 3.2 board running a C program does all the computation and returns the position estimates at a user-selectable update rate.

The ARM microcontroller is programed to sample the four linear sensor arrays with analog-to-digital converters (ADCs), one ADC channel for each sensor array. With the high-speed sampling capability of the microcontroller's ADCs (1 Msps), the system just takes approximately 300 μs to read the data (512 data points) from the four sensors [26]. All sensors are sampled simultaneously and then read sequentially by the ADC. This also reduces the CPU time spent on sampling and focuses on signal processing. The TSL1401R sensor consists of 128 photodiodes arranged in a linear array. Light energy impinging on each photodiode generates photocurrent, which is integrated by the active integration circuitry associated with that pixel. By controlling the integration time, the exposure time of the sensor can be controlled to avoid sensor saturation (for short range) and signal amplification (for large range). In addition to LOSAs, a 9-DOF IMU was connected to measure the dynamics of the tracking system. Limited by the UART serial speed and buffering issues, the sensor sent the IR data to the PC at an average update rate of ∼115 Hz.

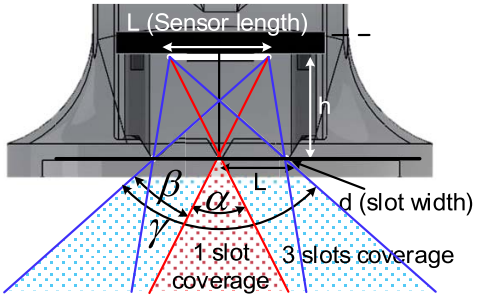

Fig. 3. Field of view of a LOSA sensor module with three slots.

TABLE I
RELATIONSHIP BETWEEN SENSOR HEIGHT AND FIELD OF VIEW

h(mm)	α (degree)	γ(degree)	d(mm)
4	90	143	0.08
8	53	112	0.11
12	**37**	**90**	**0.13**
15	24	77	0.16
25	18	51	0.20

C. Pinhole Width and Sensor Positioning

Normally, the smaller the pinhole width, the sharper the image. But if the hole is too small, diffraction effects will occur and the image will become less clear due to the wave properties of light. In addition, smooth slot edges and thin slot materials also reduce undesired diffraction. Petzval first found an easy method to calculate the optimal pinhole diameter/width [27]. The crispest image is obtained using a pinhole size determined by:

$$d = \sqrt{2h\lambda}. \tag{1}$$

where d is pinhole diameter (slot width), h is focal length (distance from slot to image plane) and λ is the wavelength of light. Further it can be shown though trigonometry (Fig. 3) that for a given slot width there exists a unique 'field of view' angle.

Due to limited field of view angle using a single slit, three optical slits were used in the sensor enclosure to increase the field of view. The distance between the adjacent slots was selected as same as the sensor length. This ensures that light passing thorough only one of the three slots creates significant image (intensity peak) on the sensor without creating any 'blind spots'. Given the IR source wavelength as $\lambda = 850$ nm and sensor length $L = 8.0$ mm the field of view angles (α, β and γ) are estimated as follows:

$$\alpha = 2\arctan\frac{L}{2h}; \quad \beta = \arctan\frac{1.5L}{h} - \frac{1}{2}\alpha; \quad \gamma \approx \alpha + 2\beta \tag{2}$$

where, L is the distance between adjacent slits (same as sensor length) and h is the sensor height from slit plane. Table I shows the dependence of field-of-view angles and slot width to sensor height. To maximize field of view (γ) to 90° with slot width as limiting factor, the sensor height was chosen as 12mm.

D. Active LED Marker Implementation

The active marker is a battery powered IR LED light (850nm) source. It consists of four independent current controlled high power infrared LEDs mounted on the surface of a PCB tetrahedron, as shown in Fig. 4. Each LED has a viewing angle of 110°, thus the tetrahedron structure enables the four LEDs to emit a signal in all directions. Each LED marker also hosts an InvenSense™ MPU9250® 9DOF IMU sensor. The marker is controlled by a marker control module equipped with ARM Cortex-M4 microcontroller (for onboard computation) and Xbee™ WiFi transceiver. The control module acquires the IMU data and integrates the angular rates to obtain attitude at an update rate of 450 Hz and finally sends the IMU and attitude data to the PC over a Wi-Fi network at an average update rate of 120 Hz. In return, it receives the LED peak values of IR intensities (received by the LOSA sensors) and position estimates from the PC.

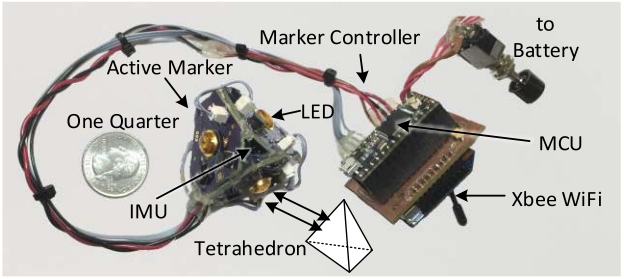

Fig. 4. IR LED light source (Active Marker) with controller.

When powered ON, the active marker enters into 'calibration mode' for 5s. During calibration mode, the marker is kept stationary to estimate 'Line of Sight' (LoS) vector (pointing towards the LOSA sensor in its frame of reference) and noise/bias onboard the gyroscope and the accelerometer (explained in subsequent sections). After the calibration period, the marker keeps track of the LoS vector (sensor direction) and the initial gravity vector by integrating rotation quaternion rates (obtained from IMU's angular rates after removing bias). The marker sends this tracked gravity and LoS vectors to the PC over Wi-Fi for attitude estimation. The marker uses the LoS vector for illuminating only that LED which faces the LOSA sensor the best. The marker computes vector dot product of the LoS vector with LED direction vectors and uses the result as expected LED intensities. Also, because of LED intensity feedback from the PC, the marker maintains the lowest LED intensity required to detect peaks reliably (20% of full-scale value), further saving power. To prevent/minimize drifting issues in gyroscope data, the marker uses fresh readings from the gyroscope to adjust gyroscope bias when the measured angular rates and accelerations are less than a preset threshold. Fig. 5 shows the complete schematics of the proposed tracking system. The marker automatically switches to Calibration mode if the IR LOSA sensor loses IR intensities for more than 3s to correct gyroscope drifting issues.

E. Estimating Peak Values and Tracking Peaks

Since the spatial resolution of the LOSA sensor is just 128 pixels, the accurate subpixel peak estimation is extremely important. Fig. 6(A) shows the raw data collected from one sensor module. It is observed that the peak width was always less than 7 pixels. Thus, a window of 7 pixels around the

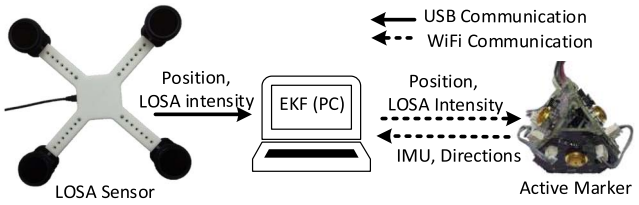

Fig. 5. Schematics of the proposed tracking system.

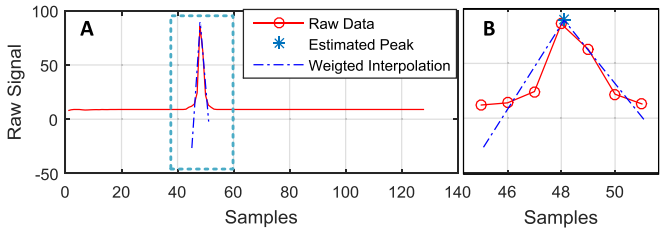

Fig. 6. Peak Detection. (A) Raw data from one sensor module. (B) Estimated peak position on zoomed peak area.

sampled peak position was used to estimate the analog peak position. To estimate the peak position, linear interpolation was used in the peak region through a weighted least squares method. First, a window of ±3 pixel (total length 7 pixels) around each signal maximum was isolated. Then, the approximate center position was estimated through a weighted mean of the indices of the pixels in the local maxima window where intensity values correspond to the weights. This approximate center was then used to divide the pixels in the window into two groups (one on the left of the peak and another on the right). Then the weighted least squares method was used to model the two groups of the points with two lines. The intersection of the lines was taken as the peak position. Fig 6(B) shows peak estimation on one sample signal.

Since the sensor module geometry ensures that light from only one slot results in a peak on the sensor, it is necessary to estimate which slot the light passed through. An algorithm was implemented to estimate and track the slot ID. It is assumed that all the peaks originate from the central slots. The microcontroller in the LOSA sensor keeps track of the location of all the detected peaks. Whenever a sudden change is detected in peak location, 128 (length of sensor in pixels) is added or subtracted from the peak positions depending on the direction of motion of the peak. This process virtually adds extra length to the optical sensor strip and hence increases the spatial range of the sensor. These peak locations are then used to estimate the actual position of the active marker (LEDs) with respect to the tracking system.

III. Data Fusion Using EKF for 3D Position and Attitude Tracking

The proposed tracking system measures position and orientation with the LOSA sensor. Since the active marker is equipped with 9-axis IMU, inertial data from the marker can be used to supplement the optical sensing data to improve position/attitude estimates. The proposed system uses Extended Kalman Filter (EKF), which is a standard approach for state estimation involving nonlinear systems [28], [29].

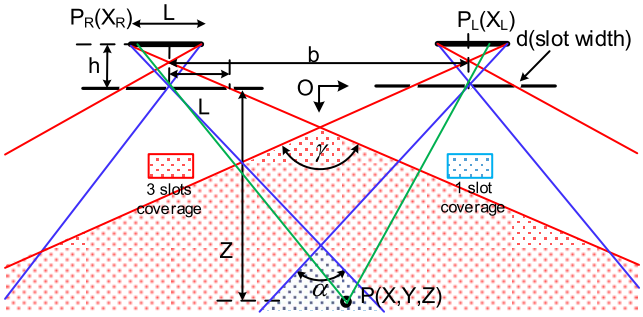

Fig. 7. LOSA sensor pair geometry for triangulation.

A. Position Measurement

Position estimates of the marker are obtained through planer stereovision. Given a pair of aligned sensors (Fig. 7), the planar (x-z plane) position of the light source may be calculated using stereo triangulation [28] as follows.

$$\frac{x_L}{h} = \frac{X - b/2}{Z}, \quad \frac{x_R}{h} = \frac{X + b/2}{Z}. \tag{3}$$

After solving (3) for X and Z, we get:

$$X = b\frac{x_L + x_R}{2(x_L - x_R)}, \quad Z = b\frac{h}{x_L - x_R}, \tag{4}$$

The position shifting between two linear sensors $x_L - x_R$ is called the disparity d. From Equation (4), depth information (Z) is uniquely determined by disparity and is inversely proportional to d. Thus, measuring object position is reduced to the problem of finding photographic mapping of the object location on each sensor (x_L and x_R) with respect to the optic center of each sensor. The optic centers of each of the four LOSA modules were estimated using a simple calibration experiment. Since the baseline is variable while the sensor's spatial resolution is limited (128 pixels), the baseline effectively defines the operational range of the system (Equation 4). As the distance between the marker and the sensor increases, the disparity decreases and approaches the noise level. This defines the upper limit of the sensing range. Increasing the baseline adds 'offset' in stereo disparity (Fig. 7) and hence raises the upper limit of the sensing range. The lower limit of the sensing range also rises with increase in the baseline due to decrease in the common coverage region (shaded region in Fig. 7). The rate of increase is greater in the upper limit than that of the lower, thus increasing the baseline increases the active range of the sensor. A similar calculation can be performed to find a (*Y,Z*) pair to obtain 3D position estimates. These position estimates are also differentiated with respect to time to obtain 3D velocity estimates.

B. Attitude Estimation

Since the IMU magnetometer is susceptible to stray magnetic fields, it cannot be used for attitude estimation. To estimate the attitude of the marker with respect to the sensor, two non-parallel vectors including LoS and Gravity are used. During the calibration mode, the marker, which is kept stationary, measures acceleration to get the gravity vector. To estimate the LoS vector, the marker lights the LEDs

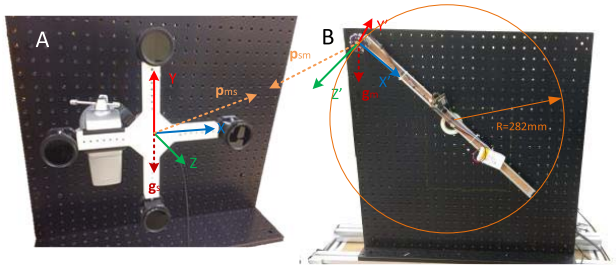

Fig. 8. Experimental Setup with coordinate frames (mounting plates are parallel and face each other). (A) tracking system. (B) Active LED marker.

(one LED at a time) in a fixed pattern and receives corresponding LED intensities from the LOSA sensors over Wi-Fi. Since the direction vectors for each LED are known in the marker's frame of reference (due to the tetrahedron shape), the LoS vector of the sensor is estimated by taking the weighted average of the LED's direction vectors. The weights of the weighted average are obtained by the IR intensities received by LOSAs during the calibration phase.

Since the LOSA cannot differentiate between simultaneous light from different LEDs, it is not possible to estimate the LoS in real-time. Therefore, the system relies on integration of angular rates for the estimation of attitude. The integration errors in attitude estimates were corrected with the help of position updates. The position and attitude estimates are linked with each other through velocity updates, as the net acceleration includes attitude information in the form of rotation matrix. The LoS vector and the gravity vector are two vectors which are known in both frames of reference (as shown in Fig. 8). These pairs of vectors can be used to obtain the attitude of the marker with respect to the sensor.

$$\begin{aligned} g_m &= R_s^m g_s; \quad g_s = [0 \quad -g \quad 0]^T, \\ \widehat{r}_{sm} &= -R_s^m \widehat{r}_{ms}; \quad r_{ms} = [r_x \; r_y \; r_z]^T, \; \widehat{r}_{ms} = \frac{r_{ms}}{\|r_{ms}\|} \end{aligned} \tag{5}$$

where g is gravity vector, R_s^m is the rotation matrix to transform coordinates from sensor to marker's frame of reference, $\widehat{r}_{ms}$ represent the direction vector, and r_{ms} is the position estimate of the marker. The subscripts s and m denote *sensor* and *marker* frame of reference, respectively. R_s^m can also be written in terms of unit quaternion vector $\boldsymbol{Q} = [\, q_0 \; q_1 \; q_2 \; q_3 \,]^T$ as shown in (6).

$$R_s^m(\boldsymbol{Q}) = \begin{bmatrix} 1-2(q_1^2+q_2^2) & 2(q_0q_1+q_3q_2) & 2(q_0q_2-q_3q_1) \\ 2(q_0q_1-q_3q_2) & 1-2(q_0^2+q_2^2) & 2(q_1q_2+q_3q_0) \\ 2(q_0q_2+q_3q_1) & 2(q_1q_2+q_3q_0) & 1-2(q_0^2+q_1^2) \end{bmatrix} \tag{6}$$

where, $\boldsymbol{Q}$ represent the rotation (attitude) of the marker with respect to the sensor's frame of reference. Since $\widehat{r}_{sm}$ and g_m are known from the measurements taken during the calibration mode of the markers, the rotation matrix R_s^m can be obtained by solving (5). This rotation matrix can be used to calculate $\boldsymbol{Q}$ and subsequently the Euler angles to represent attitude in 'human readable' format. The rotation quaternion $\boldsymbol{Q}$ is updated by integrating angular rates on the microcontroller in order to obtain the attitude of the marker in real-time.

C. System Modeling

As mentioned above, EKF was used to fuse IMU and optical tracking data to obtain position and attitude estimates of the active marker with respect to the LOSA sensor. Initial position and attitude of the active tracker were obtained from the LOSA sensor at system initialization. In addition to these position and attitude estimates, the marker's real-time angular rates and linear acceleration are used to further improve tracking accuracy with the help of EKF. The system has been modelled in a similar way as proposed in [29].

The state vector for the EKF composed of position, velocity and rotation (unit) quaternion estimates the marker with respect to the sensor's frame of reference. To fuse attitude information with linear position/velocity information, linear acceleration measurement from the marker's IMU was used. The acceleration measured by the IMU on the marker consists of three components, including gravity, rotational acceleration, and linear acceleration.

$$\overrightarrow{\boldsymbol{acc}}^s = R_m^s(\boldsymbol{Q}) \begin{bmatrix} a_{X'}^m - \omega_{Y'}^m V_{Z'}^m + \omega_{Z'}^m V_{Y'}^m \\ a_{Y'}^m + \omega_{X'}^m V_{Z'}^m - \omega_{Z'}^m V_{X'}^m \\ a_{Z'}^m - \omega_{X'}^m V_{Y'}^m + \omega_{Z'}^m V_{Y'}^m \end{bmatrix} - \begin{bmatrix} 0 \\ -g \\ 0 \end{bmatrix} \tag{7}$$

Here $\overrightarrow{\boldsymbol{acc}}^s$ represents the linear acceleration vector experienced by the marker in the tracker's frame of reference whereas $a_{X',Y',Z'}^m$ and $\omega_{X',Y',Z'}^m$ represent the IMU readings (accelerometer and gyroscope, respectively) for the marker in the marker's frame of reference. The state estimates were updated by integrating the measurements and the previous state estimates as follows:

$$\begin{bmatrix} \vec{p}_{k+1}^{\,s} \\ \vec{v}_{k+1}^{\,s} \\ \boldsymbol{Q}_{k+1} \end{bmatrix} = \begin{bmatrix} \vec{p}_k^{\,s} + \vec{v}_k^{\,s}.\delta t \\ \vec{v}_k^{\,s} + R_m^s \left(\vec{a}_k^{\,m} - R_s^m g_s - \vec{\omega}_k^{\,m} \times R_s^m \vec{v}_k^{\,s} \right).\delta t \\ \boldsymbol{Q}_k + \dot{\boldsymbol{Q}}_k(\boldsymbol{Q}_k, \vec{\omega}_k^{\,s}).\delta t \end{bmatrix} \tag{8}$$

The quaternion rate was obtained from measured angular rates as follows [30]:

$$\dot{\boldsymbol{Q}}(\boldsymbol{Q}, \vec{\omega}) = 0.5 \begin{bmatrix} q_0 & q_3 & -q_2 & q_1 \\ q_1 & q_2 & q_3 & -q_0 \\ q_2 & -q_1 & q_0 & q_3 \\ q_3 & -q_0 & -q_1 & -q_2 \end{bmatrix} \begin{bmatrix} 0 \\ \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \tag{9}$$

Since only linear velocity and position are the only observable system variables, the following observation model was used.

$$\begin{bmatrix} \vec{p} \\ \vec{v} \end{bmatrix} = \begin{bmatrix} I_3 & 0 & 0 \\ 0 & I_3 & 0 \end{bmatrix} \begin{bmatrix} \vec{p}_k^{\,s} \\ \vec{v}_k^{\,s} \\ \boldsymbol{Q}_k^s \end{bmatrix} \tag{10}$$

where I_3 represents an identity matrix of size 3.

D. EKF Implementation

As the Cortex M4 microprocessor is not powerful enough to run EKF in real time, EKF was implemented on a PC using NI LabVIEW®. In general, EKF can be used to solve for any non-linear stochastic process of the form:

$$\left.\begin{aligned} \boldsymbol{x}_k &= f(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k) + \boldsymbol{w}_k \\ \boldsymbol{y}_k &= h(\boldsymbol{x}_k) + \boldsymbol{v}_k \end{aligned}\right\} \boldsymbol{w}_k \in N\left(\vec{\boldsymbol{0}}, \boldsymbol{Q}_k\right), \quad \boldsymbol{v}_k \in N\left(\vec{\boldsymbol{0}}, \boldsymbol{R}_k\right) \tag{11}$$

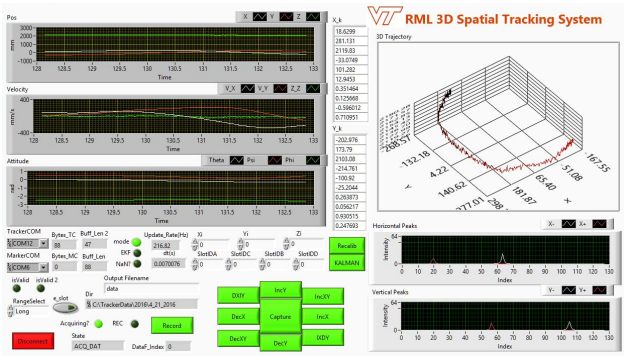

Fig. 9. LabVIEW GUI for the LOSA tracking system.

where $\boldsymbol{x}_k$ is the state vector, $\boldsymbol{u}_k$ is the control input vector and $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are the process and measurement noise vectors, respectively. The noise vectors $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are assumed to follow multivariate Gaussian distribution with $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ as covariance matrices. The state and measurement models (11) are linearized using Jacobian Matrices $\boldsymbol{F}$ and $\boldsymbol{H}$.

$$\boldsymbol{F}_k = \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\hat{\boldsymbol{x}}_k,\boldsymbol{u}_k} , \boldsymbol{H}_k = \left.\frac{\partial h}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\hat{\boldsymbol{x}}_{\mathbf{k}}} \tag{12}$$

During the 'prediction' stage of EKF, state estimates ($\hat{\boldsymbol{x}}$) and covariance matrix for state estimates ($\boldsymbol{P}$) are predicted as follows:

$$\begin{aligned} \hat{\boldsymbol{x}}_{k|k-1} &= f\left(\boldsymbol{x}_{k-1|k-1}, \boldsymbol{u}_k\right) \\ \boldsymbol{P}_{k|k-1} &= \boldsymbol{F}_k \boldsymbol{P}_{k-1|k-1} \boldsymbol{F}_k^T + \boldsymbol{Q}_k \end{aligned} \tag{13}$$

After the prediction stage, these estimates are corrected on the basis of measurements during the 'update' stage of the EKF as follows:

$$\begin{aligned} \tilde{\boldsymbol{y}} &= \boldsymbol{y}_k - h\left(\hat{\boldsymbol{x}}_{k|k-1}\right) \\ \boldsymbol{S}_k &= \boldsymbol{H}_k \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T + \boldsymbol{R}_k \\ \boldsymbol{K}_k &= \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T \boldsymbol{S}_k^{-1} \\ \hat{\boldsymbol{x}}_{k|k} &= \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \tilde{\boldsymbol{y}} \\ \boldsymbol{P}_{k|k} &= \left(\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k\right) \boldsymbol{P}_{k|k-1} \end{aligned} \tag{14}$$

where $\tilde{\boldsymbol{y}}$, $\boldsymbol{S}_k$, $\boldsymbol{K}_k$ and $\boldsymbol{I}$ represent measurement residuals, residual covariance matrix, Kalman gain, and identity matrix, respectively.

IV. Experiments and Evaluation

A. Demonstration Experiment

To acquire, process and record tracking data, a LabVIEW™ based GUI was created (Fig. 9). The GUI, running on a notebook powered by Intel core i7™ processor, was able to deliver state estimates at an update rate of more than 250Hz. The GUI also plotted the 3D trajectory of the marker at an update rate of 20Hz. Fig. 8 shows the experimental setup. For evaluating the positional accuracy of the sensor, the tracking system (sensor) was mounted rigidly on a mechanical breadboard (mounted on a guiding rail). The active marker was connected to a rod hinged on another mechanical breadboard (parallel to the tracking system plane) mounted on the same guiding rail. This setup constrained the motion of the marker on a circle in a plane parallel to the sensor plane. The maker was moved on a circular trajectory of radius 282mm from distances varying from 2m to 5.5m and position estimates were recorded. Since the position and radius of the circular trajectory is known from measurements (with accuracy ±10mm and ±1mm respectively), the measured position estimates were compared with the expected circular trajectory. To assess the positional accuracy of the tracking system, the system was tested indoors at different distances in extended range mode (400mm baseline). The extended range mode was deliberately selected because long range indoor tracking is more critical in most applications. Fig. 10 shows position estimates for the marker when rotated with the rod as shown in the experimental setup.

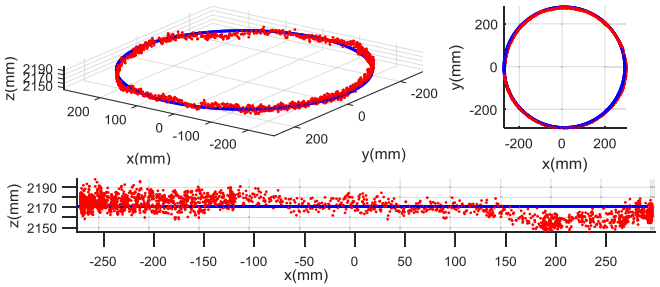

Fig. 10. Position estimates of the marker (red dots) compared with the actual trajectory (blue).

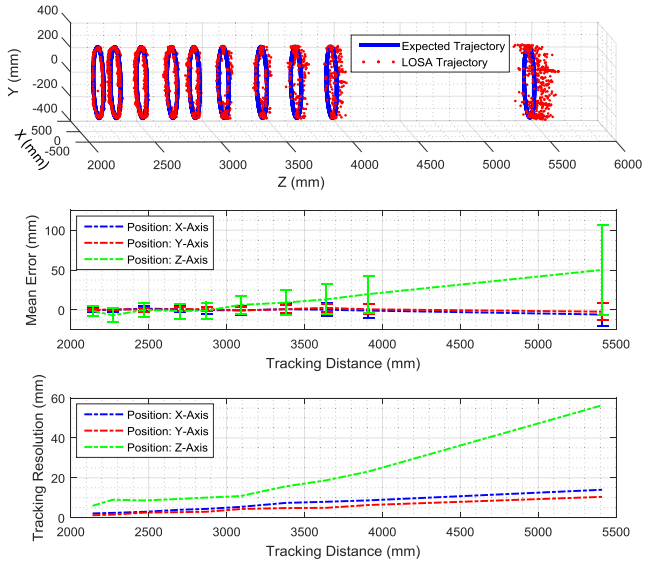

Fig. 11. Positional accuracy and resolution of the LOSA sensor at different operating ranges.

B. System Performance

Positional resolution is essentially determined by the 'signal to noise ratio' (SNR) of the sensor data. Since the signal (stereovision disparity) decreases with distance, and the sensor noise remains constant, the SNR decreases with distance and so does the resolution. The sensor resolution was defined as 'one standard deviation' (1σ) in the sensor reading. Fig. 11 compares measured position estimates from the sensor with expected trajectory and shows measured positional error in the form of mean and standard deviation for the same trajectories.

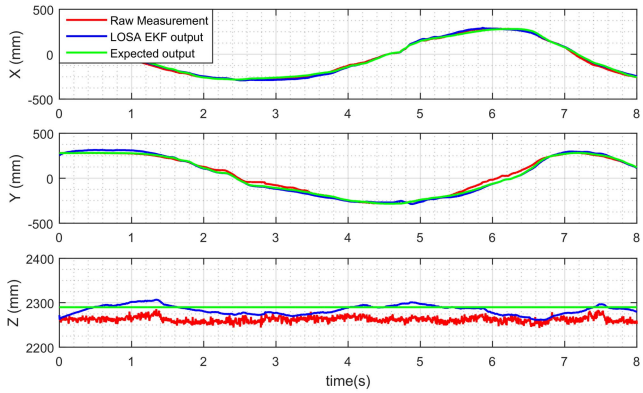

Fig. 12. Comparison of EKF output with raw positional estimates.

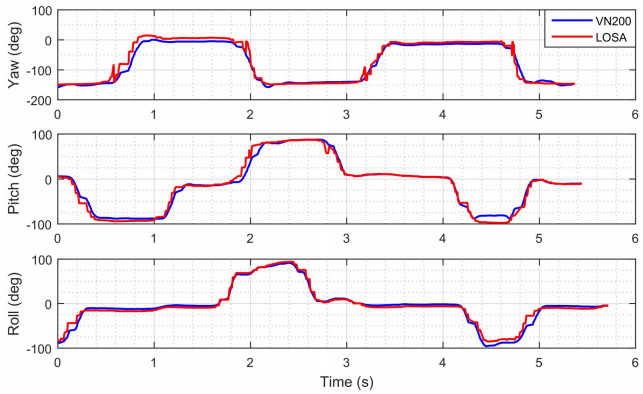

Fig. 13. Dynamic orientation measurement with VN200 INS and proposed LOSA tracking system.

It also shows the sensor resolution at different operating ranges. It was observed that the mean error (accuracy) in position estimates varied from 1.18mm to 50.56mm with tracking distance varying from 2.1m to 5.4m, respectively. EKF implementation proved to be effective in removing noise and improving position estimates. Fig. 12 compares EKF position estimates and raw position measurements with the expected values.

To assess the accuracy of the orientation estimates from the proposed sensor, Euler angles extracted in post-processing from the proposed sensor were compared with commercially available VN200 INS/AHRS sensor [31]. For this experiment, the Active marker was rigidly attached to the VN200 sensor to align measurement axes, and the assembly was rotated ($\pm 90°$) along the three axes with respect to the initial orientation. Fig. 13 shows yaw, pitch and roll estimates from the proposed LOSA sensor and VN200. Attitude estimates from the proposed sensor closely matched with those from the VN200 sensor.

C. Power Consumption

The power consumption of the four sensors (and IMU) and the controller is about 32 mA (3mA $\times$ 4 +20mA) and 20 mA at 5 V separately (around 260 mW). The tracking system was powered though the USB connection. Power consumption by the active marker is greatly reduced through optimization of the LED brightness. The marker automatically adjusts the LED intensities to light up only one LED (facing the tracker) and maintains a minimum intensity required to detect peaks reliably in the sensor. The active marker consumes about 145mA and 260mA current at 7.4V during indoor operations at 2m and 5m range, respectively. For outdoor operations, this power can be raised up to 500mA.

V. Conclusions and Future Work

A novel motion tracking technology that uses linear optical sensor array was introduced. The proposed sensor makes novel use of MEMS IMU and wireless transceivers with stereo vision to extract 3-axis attitude estimates without relying on magnetometers (which are susceptible to stray electromagnetic interferences). This system enables high speed, high precision and low-cost 3D tracking for a wide range of applications. The detailed hardware design and performance evaluation experiments of this approach are also included.

A. System Limations

Assembly Issues. The cover box is built using a 3D printer platform, which has an accuracy of 0.254 mm. Although all sensors were calibrated with the matrix's scaling parameters, those error sources are difficult to fix. Sometimes, environmental impact or mishandling of the sensor can introduce errors in the sensor geometry in the form of changes in the tilt angles/baseline. Such error can be rectified by calibrating the sensor pairs with LED array (with 3 or more LEDs mounted at known interval on a straight rod). Such array then can be used to find extrinsic parameters of the stereo rig.

Working space. Though the 3-slots method greatly increase the viewing angle of the working space, the distance between the LED and the sensor is still limited by a number of factors including the brightness and projection angle of the LED light. To increase the tracking distance, bright light source and larger detection area sensors are required. The experimental results demonstrate that the LOSA system can be used for position and attitude tracking within the spatial range of many applications.

Multiple sensors scenario. The workspace can also be increased by using multiple sensor arrays with a single marker. Since the tracking system is based on a Wi-Fi communication network, it is easy to add/remove sensor nodes from the network. This application would require the active marker to track multiple LoS vectors corresponding to each sensor array. For synchronization between multiple sensors, the active marker can be made to blink with fixed frequency for the synchronization of the Micro Processor Unit clock.

Multi light source scenarios. In our current design, the sensor unit is only able to sense one infrared LED light source, restricting its application to 3D position tracking. Time multiplexing, adding more sensor modules, or motion memorization and analysis could resolve this issue and is the focus of future work.

B. Future Work

In future work, we envision extending the LOSA tracking system in several ways. Firstly, the four sensor modules will be arranged onto one custom-built, planar mount with tight tolerances (<0.5 mm). In addition, the cover box will be

made of aluminum metal box with much higher accuracy (<0.02 mm). Secondly, using higher resolution linear optical sensors such as TSL1402R, which has 256 pixels, will also improve tracking accuracy. Also, the aperture slits may be replaced by cylindrical lenses to increase the range of the sensor by decreasing LED power requirements. Finally, an attempt will be made to reduce the size of the LOSA sensor unit by half in order to enable integration into a wider range of applications.

REFERENCES

[1] G. Welch and E. Foxlin, "Motion tracking: No silver bullet, but a respectable arsenal," *IEEE Comput. Graph. Appl.*, vol. 22, no. 6, pp. 24–38, Dec. 2002.

[2] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 37, no. 6, pp. 1067–1080, Nov. 2007.

[3] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 13–32, Mar. 2009.

[4] J. P. Rolland, Y. Baillot, and A. A. Goon, "A survey of tracking technology for virtual environments," in *Proc. Fundam. Wearable Comput. Augmented Reality*, 2001, pp. 69–112.

[5] K. Murakami *et al.*, "Position tracking system of everyday objects in an everyday environment," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, Oct. 2010, pp. 3712–3718.

[6] S. Song, H. Ren, and H. Yu, "An improved magnetic tracking method using rotating uniaxial coil with sparse points and closed form analytic solution," *IEEE Sensors J.*, vol. 14, no. 10, pp. 3585–3592, Oct. 2014.

[7] P. W. Schönle, K. Gräbe, P. Wenig, J. Höhne, J. Schrader, and B. Conrad, "Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract," *Brain Lang.*, vol. 31, no. 1, pp. 26–35, May 1987.

[8] H. Chao, L. Mao, S. Shuang, Y. Wan'an, Z. Rui, and M. Q. H. Meng, "A cubic 3-axis magnetic sensor array for wirelessly tracking magnet position and orientation," *IEEE Sensors J.*, vol. 10, no. 5, pp. 903–913, May 2010.

[9] Z. Dong, U. C. Wejinya, and W. J. Li, "An optical-tracking calibration method for MEMS-based digital writing instrument," *IEEE Sensors J.*, vol. 10, no. 10, pp. 1543–1551, Oct. 2010.

[10] R. Zhang, F. Höflinger, and L. Reindl, "Inertial sensor based indoor localization and monitoring system for emergency responders," *IEEE Sensors J.*, vol. 13, no. 2, pp. 838–848, Feb. 2013.

[11] G. V. Zàruba, M. Huber, F. A. Kamangar, and I. Chlamtac, "Indoor location tracking using RSSI readings from a single Wi-Fi access point," *Wireless Netw.*, vol. 13, no. 2, pp. 221–235, Apr. 2007.

[12] X. Luo, W. J. O'Brien and C. L. Julien, "Comparative evaluation of received signal-strength index (RSSI) based indoor localization techniques for construction jobsite," *Adv. Eng. Informat.*, vol. 25, no. 2, pp. 335–363, Apr. 2011.

[13] A. Erol, G. Bebis, M. Nicolescu, R. Boyle, and X. Twombly, "Vision-based hand pose estimation: A review," *Comput. Vis. Image Understand.*, vol. 108, nos. 1–2, pp. 52–73, Oct./Nov. 2007.

[14] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[15] C. D. Metcalf, S. V. Notley, P. H. Chappell, J. V. Burridge, and V. T. Yule, "Validation and application of a computational model for wrist and hand movements using surface markers," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 3, pp. 1199–1210, Mar. 2008.

[16] C. Celozzi, G. Paravati, A. Sanna, and F. Lamberti, "A 6-DOF ARTag-based tracking system," *IEEE Trans. Consum. Electron.*, vol. 56, no. 1, pp. 203–210, Feb. 2010.

[17] H. Woltring, "New possibilities for human motion studies by real-time light spot position measurement," *Biotelemetry*, vol. 1, no. 3, pp. 132–146, 1974.

[18] J. MacCormick. *How Does the Kinect Work?*, accessed on Sep. 15, 2015. [Online]. Available: http://users.dickinson.edu/~jmac/selected-talks/kinect.pdf

[19] V. Milanović and A. Kasturi, "Real-time 3D tracking," *Opt. Photon.*, vol. 8, no. 4, pp. 55–59, 2013.

[20] R. Xiao, C. Harrison, K. D. D. Willis, S. E. Hudson, and I. Poupyrev, "Lumitrack: Low cost, high precision and high speed tracking with projected m-sequences," in *Proc. 26th Annu. ACM Symp. User Interface Softw. Technol. (UIST)*, New York, NY, USA, Oct., 2013, pp. 3–12.

[21] A. Mohan, G. Woo, S. Hiura, Q. Smithwick, and R. Raskar, "Bokode: Imperceptible visual tags for camera-based interaction from a distance," *ACM Trans. Graph. SIGGRAPH*, vol. 28, no. 3, Jul. 2009, Art. no. 98.

[22] Q. Wen and J. Wu, "Linear CCD based optical tracking using stereo correspondence algorithm," in *Proc. Int. Conf. Artif. Intell. Comput. Intell. (AICI)*, Nov. 2009, pp. 422–425.

[23] 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. Exp.*, vol. 23, no. 3, pp. 3412–3426, Feb. 2015.

[24] Y. Ruichek, "A hierarchical neural stereo matching approach for real-time obstacle detection using linear cameras," in *Proc. IEEE Intell. Transp. Syst.*, Oct. 2003, pp. 299–304.

[25] J. C. Burie and J.-G. Postaire, "Enhancement of the road safety with a stereo vision system based on linear cameras," in *Proc. IEEE Intell. Veh. Symp.*, Sep. 1996, pp. 147–152.

[26] *PJRC.COM, LLC*, accessed on Jun. 28, 2015. [Online]. Available: https://www.pjrc.com/teensy/index.html

[27] L. Rayleigh, "Lord Rayleigh on pin-hole photography," *London, Edinburg Dublin Philos. Mag. J. Sci.*, vol. 31, no. 5, pp. 87–99, Feb. 1891.

[28] X. Yun and E. R. Bachmann, "Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking," *IEEE Trans. Robot.*, vol. 22, no. 6, pp. 1216–1227, Dec. 2006.

[29] H. Zhao and Z. Wang, "Motion measurement using inertial sensors, ultrasonic sensors, and magnetometers with extended Kalman filter for data fusion," *IEEE Sensors J.*, vol. 12, no. 5, pp. 943–953, May 2012.

[30] VectorNav. *Application Note: Quaternion Math*, accessed on Mar. 23, 2016. [Online]. Available: http://www.vectornav.com/docs/default-source/documentation/vn-100-documentation/AN002.pdf?sfvrsn=10

[31] VectorNav. *VN200 INS Specifications*, accessed on Apr. 10, 2016. [Online]. Available: http://www.vectornav.com/products/vn200-rugged/specifications

Anil Kumar (S'15) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology (IIT) Roorkee in 2010. He is currently pursuing the Ph.D. degree at the Virginia Polytechnic Institute and State University (Virginia Tech under the supervision of Prof. P. Ben-Tzvi. He was with the Biomedical Instrumentation Lab, IIT Roorkee, as a Research Fellow. His research interests include machine learning, signal and image processing, autonomous robotics, computer vision, and computer-aided instrumentation.

Pinhas Ben-Tzvi (S'02–M'08–SM'12) received the B.S. (*summa cum laude*) degree in mechanical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, and the M.S. and Ph.D. degrees in mechanical engineering from the University of Toronto, Toronto, Canada. He is currently an Associate Professor of Mechanical Engineering and Electrical and Computer Engineering, and the Founding Director of the Robotics and Mechatronics Laboratory , the Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, USA. He joined the University of Toronto in 2002. He was a Research and Development Engineer with General Electric Medical Systems Company, where he was involved in development of medical diagnostic robotic and mechatronic systems. His current research interests include robotics and intelligent autonomous systems, mechatronics, human–robot interactions, dynamic systems and control, mechanism design and system integration, and novel sensing and actuation. Application areas are varied and range from search and rescue on rough terrain to medical diagnostics, surgery, and therapy. He is a recipient of the 2013 GW SEAS Outstanding Young Researcher Award and the GW SEAS Outstanding Young Teacher Award, and several other honors and awards. He is a Technical Editor of the IEEE/ASME TRANSACTIONS ON MECHATRONICS and an Associate Editor for the *International Journal of Control, Automation and Systems*. He is a member of the American Society of Mechanical Engineers.