Point Tracking with Lensless Smart Sensors

Lizy Abraham, Andrea Urru, Mariusz P. Wilk, Salvatore Tedesco, Michael Walsh, Brendan O'Flynn
Wireless Sensor Networks Group, Micro & Nano Systems, Tyndall National Institute
University College Cork, Cork, Ireland

***Abstract*—This paper presents the applicability of a novel Lensless Smart Sensor (LSS) developed by Rambus, Inc. in 3D positioning and tracking.The unique diffraction pattern attached to the sensor enables more precise position tracking than possible with lenses by capturing more information about the scene. In this work, the sensor characteristics is assessed and accuracy analysis is accomplished for the single point tracking scenario.**

Keywords—Rambus LSS; Image Reconstruction; Point Tracking

I. INTRODUCTION

Real-time point positioning and tracking in 3D has been an important research topic in the field of computer vision, pattern recognition, and human–computer interaction, also involving object identification and gesture recognition. The development of depth sensors has enhanced various point tracking approaches and applications, which were severely limited in 2D domain with conventional vision-based cameras. The recent introduction of consumer depth sensors, e.g., Kinect, leap motion, and Time-of-Flight (ToF) sensors [1-3], has greatly promoted the research on 3D point tracking. Other than depth sensors, Inertial Measurement Units (IMUs) are also widely used for tracking applications [4]. IMUs are typically adopted for measuring motion and movements, however, they may suffer from drifts and cannot provide the absolute reference of a point in the world coordinate system. Thus, fusion of inertial and depth sensors are being developed for such applications [5]. Nowadays, all the major consumer developers, i.e., Google, Facebook, etc., are developing computer interfacing/gesture control technologies and investing in the area. The most basic requirement for a product is to track a single point accurately and reliably.

This work is focused on researching new potential IoT applications for the novel Lensless Smart Sensor (LSS) developed by Rambus, Inc.[6,7]. The paper describes an improved method for 3D positioning and tracking of a point source of light using the developed LSS, in comparison to the results presented in [8]. This research focused on point tracking using the LSS will eventually lead to multipoint tracking and, thereby, object tracking and gesture recognition provided that the points are tracked down to mm-level accuracy, which is the main target for this work. This is achieved by precise characterization of sensors, tracking points down to sub-pixel level, and developing an improved compensation function which provides better fitting in every point of the application area.The paper is organized as follows. The operational behavior of Rambus LSS is given in section II. The methodology related to image reconstruction, 3D point tracking and distortion compensation are discussed in Section III. Result analysis and performance evaluation are described in Section IV, while conclusions are finally drawn in Section V.

This work was supported by CONNECT, the Science Foundation Ireland (SFI) research centre for Future Networks and Communications and Rambus, Inc., USA.

E-mails: name_author.surname_author@tyndall.ie

II. SYSTEM FUNDAMENTALS

In the adopted CMOS-based imaging sensor, the lens used in conventional cameras is replaced by a low-cost ultra-miniaturized diffraction optic attached directly to the image sensor array. Light passes through the LSS gratings and creates predictable light patterns on the sensor based on the characteristics of the diffraction optics. This data-rich 'blob' is then processed by advanced algorithms to provide meaningful data, as detailed in [8].Fig. 1 shows an example of the pattern obtained for a point source of light and the corresponding reconstructed point. The net result is a smart sensor with a smaller form factor, higher power efficiency and precision, and lower cost compared to traditional visual sensors. As images and sensor decisions are not captured directly, privacy and minimal data transfer are ensured.

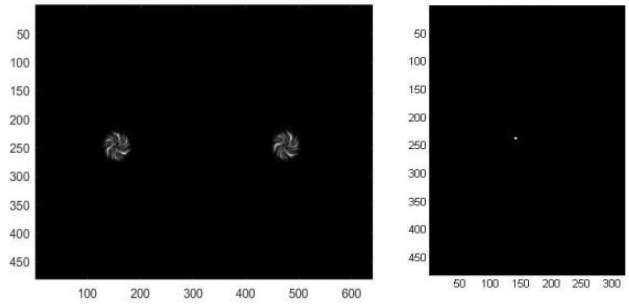

Fig. 1. (a) Point Source seen by a LSS (b) Reconstructed Point

III. METHODOLOGY

Based on the novel Rambus LSS, the sensor technology is assessed to establish the opportunities for single point live tracking. For tracking 3D positions of a point source without providing any manual inputs, at least two sensors are required. As a first step, data acquisition was carried out using individual sensors separately. The main objective was to examine the characteristics of each sensors and estimate the maximum workable area of the system. Then, a blind test, in which the 3D positions of the point source is tracked in real-time without providing any inputs, using dual sensors is conducted and the results are compared.

A. System Components

The selection of an appropriate light source is crucial to ensure the best possible accuracy. An inexpensive infrared LED – filter pair for ambient light isolation was used in the test. The LED was selected based on power consumption,

reduced diameter and yet wide viewing angles, to ensure that the reconstructed points are sharper and distinguishable even at higher viewing angles. The choice of the infrared filter is based on the transmission characteristics matching with the LED wavelength.

B. Image Reconstruction

The reconstruction of a point source from the raw images having two spirals is carried out as shown in [8].To improve accuracy further, the points are tracked down to sub-pixel level. For this purpose, three different best possible techniques [9,10], namely Ratio based method (1-2), Gaussian Estimator (3) and Simplified Linear Interpolation (4), are tested and examined. If $f(x,y)$ is the centre point at pixel level, $f(x+1,y)$and$f(x-1,y)$being the two surrounding pixels along x-direction, the subpixel offset dx for each method is calculated as:

$$dx = \frac{px(2)}{px(1)} \tag{1}$$

$$\begin{bmatrix} px(0) & px(1) & px(2) \end{bmatrix}^T = \begin{bmatrix} -0.5 & 1 & -0.5 \\ -0.25 & 0 & 0.25 \\ 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} f(x-1,y) \\ f(x,y) \\ f(x+1,y) \end{bmatrix} \tag{2}$$

$$dx = \frac{0.5 \times [f(x-1,y) - \log(f(x+1,y))]}{[\log(f(x-1,y)) - 2 \times \log(f(x,y)) + \log(f(x+1,y))]} \tag{3}$$

$$dx = \frac{f(x+1,y) - f(x-1,y)}{f(x,y)} \tag{4}$$

Likewise, it is possible to calculate the y-direction offset dy using $f(x,y+1)$and$f(x,y-1)$.Considering accuracy in terms of root mean square error (RMSE) and latency in terms of execution time for the reconstruction of a single video frame, sub-pixel point detection based on Gaussian estimator is chosen for this work as it has the best accuracy with an average execution speed (Table I).

TABLE I. SUBPIXEL POINT DETECTORS: RMSE AND EXECUTION TIME

Sub-pixel Point Detectors	RMSE (cm)	Execution time (s)
Ratio based Method	0.83	0.41
Gaussian Estimator	0.79	0.39
Simplified Linear Interpolation	1.21	0.36

C. 3D Point Tracking

The 3D position of a point source of light (Fig.2) is calculated similarly to [8].The longitudinal distance Z can be calculated using law of trigonometry and Snell's law of refraction. The ranging algorithm is combined for both LSS simultaneously to implement 3D single point tracking in real-time. The light source is moved from -40degto +40degfrom the centre axis and error in both X and Z axis are estimated. From the performance of the system in the dual sensor case, it is noticed that the error within the region between both sensors is lower and it increases in a symmetrical manner when the field of view (FoV) is increased.

D. Distortion Compensation

From the sensor characteristics, it is realized that the overall accuracy of the ranging algorithm can be improved by correcting the angles between focal points of the right/left sensors and the points where the photons and the LSS are in contact, via a polynomial fit function. The distortion compensation is modelled by collecting a series of images of the LED moved on a line parallel to the sensor plane at a fixed Z distance and by placing the LED at different random positions varying Z distance. For the first set of data, the longitudinal distance is 40cm from sensor plane, and data are collected from 61 positions, spaced 1cm, along the X-axis in order to cover the whole FoV. For the second set of data, the longitudinal distance Z is varied to cover the entire application area of 1m, for overall 60 random positions. For each point, 5 raw frames were acquired by both sensors. A polynomial fit function is then obtained by calculating the coefficients of the fitting curve which minimizes the RMS error. With this new approach, accuracy, compared to results in [8], is improved both laterally and longitudinally.

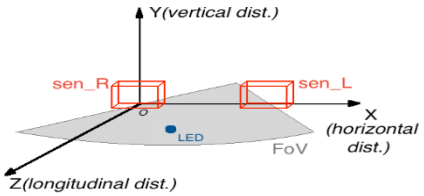

Fig. 2. System Overview (sen_L – Left sensor, sen_R – Right sensor)

IV. RESULTS AND DISCUSSION

As an initial analysis, a non-blind tracking is conducted for each sensor in which Z is given as an additional input and examined the variation on X-axis with respect to Z-axis when the light source is moved longitudinally from 40 to 100 cm. Next, Z is calculated using both sensors and the variation on the X-axis is examined in the same way (Fig. 3).

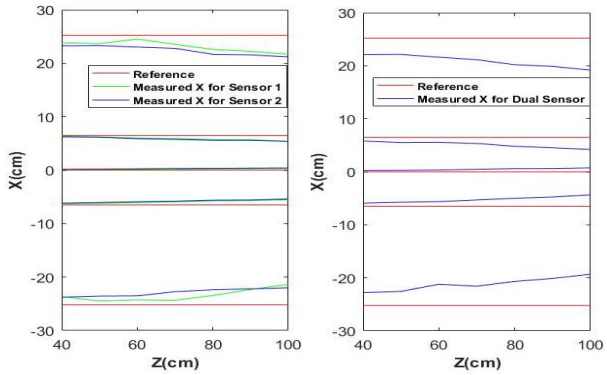

Fig. 3. Measured X vs Z (a) for eachSensors (non-blind tracking) (b) using both Sensors (blind tracking)

From Fig.3a, it is noticed that, when the light source is perfectly aligned with respect to the sensor optical axis (e.g. when X=0), the tracked path is nearly aligning with the reference, even when the light source is moved away from sensor plane up to 1m.Table II shows the RMS error for three different cases: along the optical axis of the sensors, and on the external sides of the system, with the angle α, defining the FoV, equal to ±40deg. For all three cases, the total RMS error

is then calculated by varying the longitudinal distance from 40 to 100cmwith 10cm spacing. For the dual sensor case with unknown Z, Fig.3bshows the variation over the X-axis and Fig.4a shows the variation along the Z-axis. Being a blind tracking, with no manual inputs, it is remarkable that the accuracy obtained at centre axis of the plane, and along sensor 1 and sensor 2 optical axes is approximately equal to 1cm along X-axis considering the entire 1m application area (Fig.3b). The increased error, especially at wider viewing angles and longer distances observed from fig.4a, is considerably reduced all over the application area using the compensation function. Fig.4b shows the positive effects of the distortion compensation on the overall performance of the algorithm throughout the entire area. It can be seen that the measured path of the light source is aligning with the reference line even when the light source is moved away from the sensor up to 1m showing a considerable improvement compared to [8].

TABLE II. RMS ERROR FOR EACH SENSOR

RMSE (cm)	Left (α=-40^0)	Centre (a=0^0)	Right (α=+40^0)
Sensor 1	1.011	0.194	0.864
Sensor 2	1.220	0.252	0.942

Table III shows the accuracy along X-axis in terms of RMS error with/without compensation function in five different cases: along the centre of the baseline, along the optical axis of two sensors, and along the external sides of the application area where the starting points of those lines are obtained for Z=40 cm and the angle α equal to ±40deg.For all five cases, the total RMSE is calculated in a similar way by moving the light source longitudinally. It is noticed that the compensation reduces the error to 2mm along the X-axis at center position. Also in between the two sensors, error is less than 3 mm. Even when the light source is away with a FoV of ±40deg where not enough photons are obtained in the sensor, the compensation function improves the accuracy to less than 1cm which is remarkable about the sensors.This proves that the unique diffraction pattern enables more precise position tracking than possible with a lens by capturing more information about the scene.

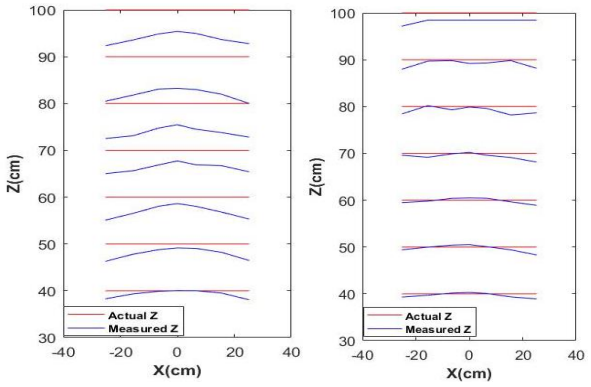

Fig. 4. Measured Vs Actual Z (a) without Compensation (b) with Compensation

Table IV illustrates the final accuracy of the single point live tracking algorithm (along Z-axis) in terms of RMS error with/without compensation function. The RMSE obtained along the Z-axis, within the FoV of 80deggiven by placing the sensors 13cm apart, is lower than 1.8cm in all five cases. Moreover, the new distortion compensation approach allows a more uniform 3D positioning along the entire FoV at the price of slightly worse results at the center, compared to the results obtained with the previous compensation used in [8].

TABLE III. RMS ERROR FOR DIFFERENT ANGLES ALONG X-AXIS (Z BETWEEN 40 AND 100CM)

RMSE along X-axis (cm)	Left (α=-40^0)	Sen_L Centre	Center (α=0deg)	Sen_R Centre	Right (α=+40^0)
with compensation	0.674	0.262	0.223	0.281	0.812
without compensation	2.068	0.919	0.613	1.024	2.814

TABLE IV. RMS ERROR FOR DIFFERENT ANGLES ALONG Z-AXIS (Z BETWEEN 40 AND 100CM)

RMSE along Z-axis (cm)	Left (α=-40^0)	Sen_L Centre	Center (α=0deg)	Sen_R Centre	Right (α=+40^0)
with compensation	1.56	0.65	0.57	0.84	1.83
without compensation	5.200	4.781	4.284	4.959	5.466
With compensation [8]	2.13	N/A	0.29	N/A	2.47

V. CONCLUSIONS

The present work proves the suitability of the Rambus LSS for tracking and ranging applications. The accuracy achieved with the new approach is promising and the technology can be easily adopted for multiple points tracking in real-time. The sensing unit will be also integrated into a wearable hand/fingers motion tracking system for infrastructure-less virtual reality applications.

ACKNOWLEDGMENT

The authors thank Evan Erickson, Patrick Gill and Mark Kellam from Rambus, Inc. for the considerable technical assistance.

REFERENCES

[1] Z. Zhang, "Microsoft Kinect sensor and its effect," IEEE MultiMedia,19, 2, pp. 4–10, Feb. 2012.

[2] F. Weichert, et al., "Analysis of the accuracy and robustness of the leap motion controller," Sensors, 13, 5, pp. 6380–6393, 2013.

[3] R. Nair et al., "A survey on time-of-flight stereo fusion," In *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*, Springer, pp. 105–127,2013.

[4] B. O'Flynn et al., "Integrated smart glove for hand motion monitoring", Int. Conf. on Sensor Technologies and Applications, pp. 45-50, 2015.

[5] K. Liu, et al., "Fusion of inertial and depth sensor data for robust hand gesture recognition," IEEE Sensors J, 14, 6, pp. 1898-1903, 2014.

[6] D.G. Stork, et al., "Optical, mathematical, and computational foundations of lensless ultra-miniature diffractive imagers and sensors", Int. J. Advances in Sys Meas, 7, 3-4, pp. 201-208, 2014.

[7] P. Gill, et al., "Lensless smart sensors: Optical and thermal sensing for the Internet of Things," IEEE Symp. VLSI Circ., 2016.

[8] L.Abraham et al., "3D ranging and tracking using lensless smart sensors", Int. Conf. Smart Systems Integration, pp. 1-8, March 2017.

[9] M.P. Wilk et al., "Sub-pixel point detection algorithm for point tracking with low-power wearable camera systems", Irish Signals and Systems Conf., ISSC, 2017.

[10] R.B. Fisher, et al., "A comparison of algorithms for subpixel peak detection. Advances in Image Processing", Multimedia and Machine Vision, Springer, pp. 385-404,1996.