

# Six-DoF Pose Estimation Using Dual-Axis Rotating Laser Sweeps Using a Probabilistic Framework

Dennis Laurijssen<sup>\*†</sup>, Steven Truijen<sup>†‡</sup>, Wim Saeys<sup>†</sup>, Walter Daems<sup>\*‡</sup>, Jan Steckel<sup>\*‡§</sup>

<sup>\*</sup>FTI CoSys Lab, University of Antwerp, Belgium

<sup>†</sup>FMH REVAKI, University of Antwerp, Belgium

<sup>‡</sup>Centre for Care Technology, University of Antwerp, Belgium

<sup>§</sup>Flanders Make Strategic Research Centre, Belgium

**Abstract**—Pose estimation systems have seen some big developments in the last two decades due to technological advances and a greater need for these systems. The industry that thrives and simultaneously popularizes these developments is the entertainment industry. One of the latest developments in home entertainment are Virtual Reality systems that lets users have an immersive experience when playing video games. In order to fully achieve this immersive experience, objects in the ‘real world’ can be used to manipulate objects in the virtual world. Therefore the ‘real world’ objects need to be located in the user’s environment to a great extent. One of these virtual reality kits utilizes an optical solution that incorporates dual-axis rotating laser sweeps on the transmitter side and photodiodes on the receiver side to achieve the positioning of these objects. Since the transmitters are stand-alone systems that can be purchased at a relatively low price, we propose to use this hardware in combination with custom low-cost receiver hardware to achieve an affordable yet precise and accurate six degrees-of-freedom human body pose estimation system.

**Index Terms**—Pose Estimation, Motion Capture, Sensor Arrays, Distributed Embedded Systems

## I. INTRODUCTION

Every digital and analog system requires some form of input to produce an output. Complex digital systems such as computers use peripherals to capture and digitize human input for performing interactions. These peripherals range from keyboards, mice, microphones and since the last few decades intricate human body pose estimation systems (more popularly known as motion capture systems) have enabled us to digitize an entire human body in order to use it as an input for computer interaction. These motion capture systems have been introduced to the broad public by the entertainment industry where they are used to produce realistic and lifelike computer generated animations in movies and video games. This technology has also sparked the creation of new input devices for controlling the actions within video games, all of the popular console gaming systems have adopted this form of input for over a decade (e.g. the Nintendo Wii Remote or the Microsoft Kinect). Recently the video games industry has introduced Virtual Reality (VR) technology to a broader public which presents users an immersive experience when playing video games. The VR experience is foremost created by the Head Mounted Display (HMD) that is worn on the head as

a sort of glasses, it features two small screens for producing a stereoscopic 3D image and is able to track the user’s head motions to adjust the viewing angle into the 3D world. In order to create a truly immersive experience the manufacturers of these VR kits have created controllers that enable the user to interact with objects in the 3D world by using gestures. This means that these controllers are actively being located in the ‘real world’ and that their positions and orientations are translated into the digital world. The HTC Vive [1] is one of the VR systems that offers this functionality. It positions and tracks both the HMD and the controllers with separate devices called HTC Lighthouses which need to be placed in the environment. The advantage of these Lighthouses is that they can work as a stand-alone unit (without needing a computer) and inherently can synchronize with other Lighthouses without the need of any user interaction.

Since the introduction of the Microsoft Kinect that can be used with a personal computer and its free Software Development Kit (SDK), it has evolved from an input device for playing video games to a low-cost device for conducting research into topics such as indoor (3D) mapping [2], [3] and human pose estimation in combination with gait analysis [4]–[10]. Although impressive advancements have been made in the research field of the latter, the Microsoft Kinect can not achieve the required accuracy and precision [11] that is required for a number of human body pose estimation or gait analysis use cases. The use cases that require highly accurate and precise human body pose estimation system currently rely on commercial products such as the camera-based ViCon measurement system or the IMU-based Xsens system [8], [12]–[15]. Although these systems boast impressive accuracy and precision, the high costs in combination with the difficulty of setting up these systems makes them unattractive and out of reach for practical use.

Like the Microsoft Kinect has been used in applications besides of its initial purpose, we propose using the HTC Light-house devices in combination with custom low-cost hardware to create an affordable six degrees-of-freedom human body pose estimation system that features the required accuracy and precision for human motion analysis (e.g. objective assessment of rehabilitation therapy of stroke patients). Whereas in previous work we devised techniques to utilize ultrasound [16]–

[18] for creating this type of human body pose estimation, the proposed system in this paper is an alternative for ultrasound transmitters and receivers and relies on the HTC Lighthouse's infrared dual-axis laser sweeps for estimating the positions (in  $X$ ,  $Y$  and  $Z$ ) of mobile receiver nodes and their orientations  $\alpha$ ,  $\beta$  and  $\gamma$  as the rotation around the three principle axes.

In the following section an overview will be given on the general system topology of the proposed system whereas the hardware implementation of the measurement prototypes and the pose estimation algorithm will be described in section III. The conducted experiments and results using the proposed measurement system will be presented in the fourth section, followed by the final section in which the conclusion and future work will be discussed.

## II. SYSTEM TOPOLOGY

Compared to our previous work for performing pose estimation [16]–[18], which proposed an acoustic solution utilizing the Time of Flight (ToF) between multiple simultaneous ultrasound transmitters and a microphone receiver array fitted on a mobile receiver node, the currently proposed measurement system is an optical solution that relies on infrared light. In the proposed measurement system two HTC Lighthouses, which we will designate as  $LH_A$  and  $LH_B$ , are placed in the environment at known positions and orientations. These two Lighthouses each feature a powerful infrared LED array and two rotating infrared lasers that can produce sweeps in both the horizontal and vertical plane. A custom mobile receiver node was designed with a photodiode array ( $PD_1$ ,  $PD_2$  and  $PD_3$ ) that is able to receive the emitted infrared light which is captured and discretized as a 1-bit digital signal in the time domain. Using these discretized signals in combination with a processing algorithm, the azimuth and elevation between the two Lighthouses and every photodiode are calculated. The azimuth and elevation measurements, relative to the Lighthouse that swept the environment, can be represented as a line between every Lighthouse and photodiode pair, as can be seen in Fig 1. Since these infrared emissions are discretized as 1-bit digital signals, this data does not contain any information regarding the distance to the point in space where the photodiodes are located. However, since the mobile receiver node's photodiode elements in the array are mounted on a custom Printed Circuit Board (PCB), the relative coordinates of the individual elements are known and by construction lie in the same plane. This information is key for estimating the position and orientation of the mobile receiver node in the environment.

In this paper two different implementations have been made for establishing the mobile receiver node, the first implementation makes use of the core PCB of our flexible hardware platform developed in earlier work [16] in combination with a photodiode array add-on hardware board which together form a single receiver node. The second implementation is built on the core hardware board's design but presents the possibility of attaching five photodiode array add-on boards. In this case we consider every add-on board as a single receiver node.

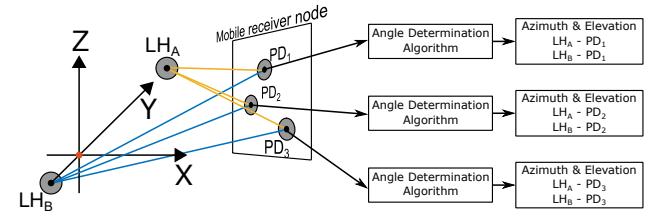


Fig. 1. System topology of the proposed measurement system, with two HTC Lighthouses (designated with  $LH_A$  and  $LH_B$ ) placed in the environment at known positions relative to the reference point in the world coordinate system. A mobile receiver node, which has been accommodated with three photodiodes  $PD_1$ ,  $PD_2$  and  $PD_3$ , in the environment will receive and capture the alternating laser sweeps as a discretized signal in the time domain for every photodiode. When these discretized signals are fed through the processing scheme it will return the azimuth and elevation between every Lighthouse and photodiode pair relative to the position of the emitting Lighthouse.

When a node is placed in the environment where the HTC Lighthouses have been installed, the emitted infrared signals will be received by each photodiode. The discretized signal that is received by a single photodiode can be seen in Figure 2 which consists of synchronization pulses and the received laser sweep signals. In Figure 3 the discretized received signals of every element of the photodiode array of the mobile node have been plotted to illustrate that every element receives the synchronization pulses simultaneously where the laser sweep will be received at a different times which results in different azimuth and elevation measurements for every photodiode.

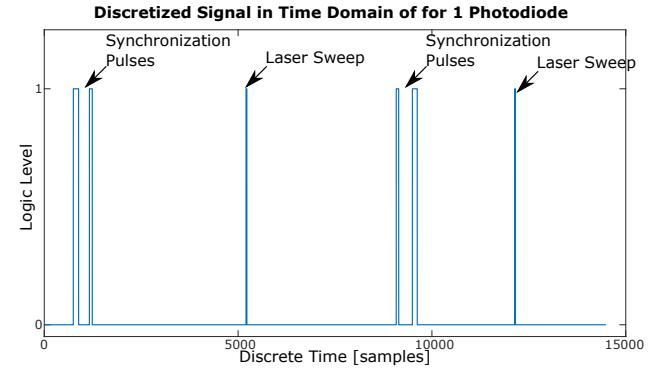


Fig. 2. The discretized received signal for a single photodiode is shown in this figure. This signal consists of synchronization pulses and both the horizontal and vertical laser sweeps which are used to determine the azimuth and elevation between the photodiode and the Lighthouse that emitted the laser sweeps. In order to determine the azimuth or elevation angle between the Lighthouse and the receiving photodiode, the  $\Delta T$  between the rising edges of the synchronization pulse and the laser sweep pulse of the corresponding Lighthouse is used.

## III. PROPOSED MEASUREMENT SYSTEM

As mentioned in the previous section, two implementations have been created for establishing the proposed measurement system. In previous work we presented a flexible hardware platform that could be used as a research tool for fast development of hardware prototypes [17] which features essential components for a multitude of applications on a main

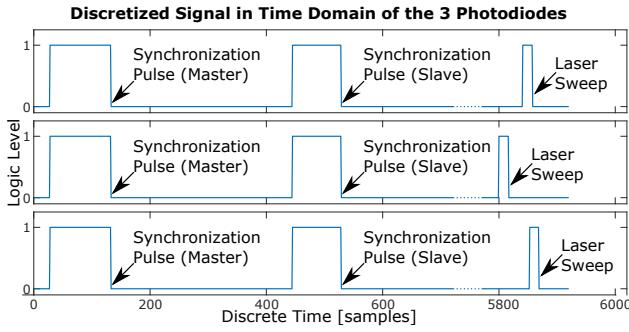


Fig. 3. This figure shows the received signals for three photo diodes on a single mobile receiver node. It can be noted that there is a noticeable difference between the arrival times of the narrower pulses, which represent the time a laser sweeps across the photodiode. These time differences enable the system to designate different azimuth and elevation values for every photodiode.

core hardware PCB that can be expanded for specific use-cases through add-on hardware boards. For this application we have therefore opted to use the core hardware board with an add-on photodiode array board as a first prototype. Figure 4 presents a schematic overview of the build-up of the flexible hardware platform with an ARM Cortex M4 microcontroller from ST Microelectronics (STM32F407VG) at its core. This 32-bit microcontroller can be clocked up to 168MHz and features 196kB SRAM, two Direct Memory Access (DMA) controllers, an integrated Floating Point Unit (FPU) and has a collection of DSP instructions. Besides these powerful attributes the microcontroller has a great number of integrated interfaces (e.g. 3x 12-bit ADC, 2x DAC, SPI, U(S)ART, SDIO, I2C, etc.) of which some have been used to connect to the other on-board peripherals such as a USB connection for easy data transfers through the FT232H IC, which bridges the USB protocol to U(S)ART, and powers the board. The integrated LSM9DS0 9-DoF Inertial Measurement Unit (IMU) can provide 3D-linear acceleration data together with 3D-angular rate data and 3D-magnetometer data that will be used in future work. For easily expanding the data storage capabilities an SD-card connector is accommodated on the core hardware board together with some headers for applying external power, debugging or connecting the add-on boards.

In order to estimate the pose of a mobile receiver node by receiving the infrared light emissions originating from the HTC Lighthouses, an add-on board was created that features a small three-element infrared photodiode array that has been arranged as an irregularly shaped triangle. For this array high speed and high sensitive SMD PIN photodiodes (Vishay VBP104S) with peak spectral sensitivity of 940nm have been chosen. For capturing the incoming infrared light with the microcontroller the diodes can not be directly connected to the microcontroller input pins, therefore the cathode of the photodiode is connected to the negative input of an operation amplifier that is configured as a trans-impedance amplifier. The output of this amplifier is passed through to

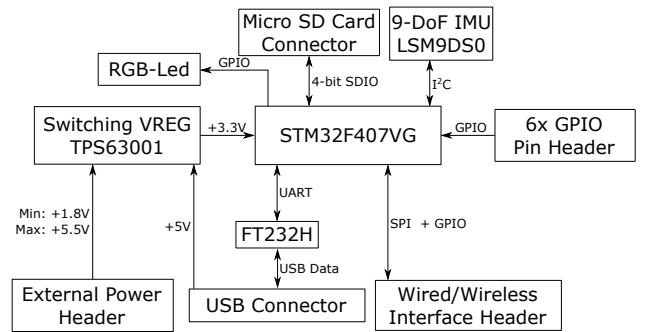


Fig. 4. Schematic overview of the core hardware board of our flexible hardware platform with the Cortex-M4 microcontroller from ST Microelectronics at its core.

a non-inverting amplifier by a high-pass filter. The output of the non-inverting amplifier will produce the discretized received signal between 0 and +3.3V and connects to a GPIO input of the microcontroller on the core hardware board. As can be seen in the schematic overview of this add-on board in Figure 5, every photodiode requires a two stage amplifier, the OPA2350 IC, which features two operational amplifiers in one package, has been used. Figure 7 shows the first implementation of the proposed mobile receiver node that comprises of an add-on board that is plugged into a core hardware board, which will be referred to as the mobile receiver node hardware stack.

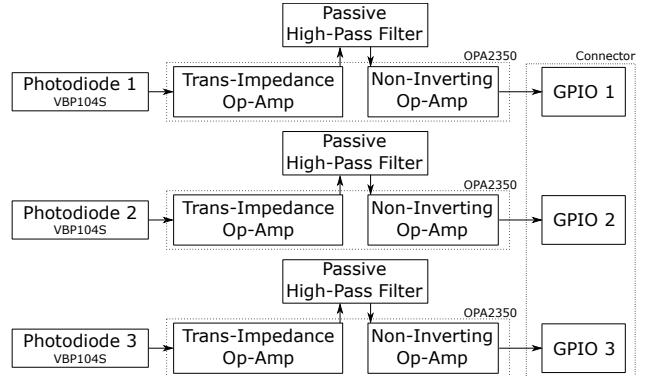


Fig. 5. Schematic overview of a photodiode receiver add-on board which features three photodiodes in combination with a trans-impedance amplifier, a high-pass filter and a non-inverting amplifier for every channel in order to connect these signals to the microcontroller board.

The hardware stack enabled us to easily perform measurements and implement the necessary software to estimate the pose of a mobile node. The goal of our research however is to establish a human body pose estimation measurement system and therefore raises the need for estimating the full pose (position and orientation) of multiple nodes simultaneously. To establish this multi-node measurement system, the aforementioned implementation of the hardware stack would require a wired or wireless synchronization scheme that entails more development work

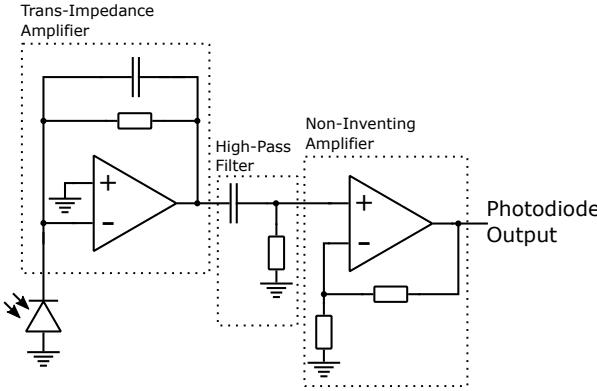


Fig. 6. Analog schematic of one channel of the a photodiode receiver add-on board that is divided into three parts: the trans-impedance amplifier, the high-pass filter and the non-inverting amplifier. The output at the second stage is the discretized photodiode signal.

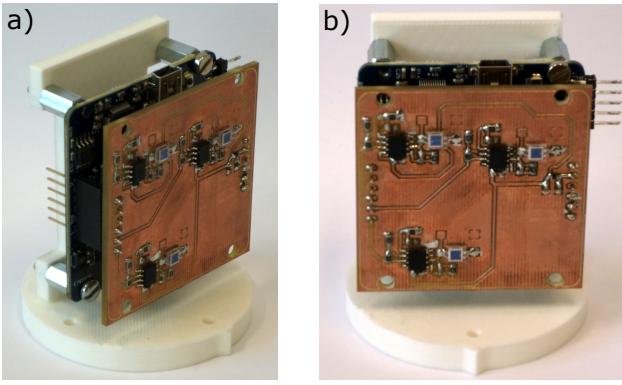


Fig. 7. a) Side view of the hardware stack system: The core hardware board is connected to a photodiode receiver board mounted on a 3D-printed fixture. b) Front view of the hardware stack system: The core hardware board is connected to a photodiode receiver board mounted on a 3D-printed fixture. The non-equilateral triangular configuration of the photodiodes can be noticed in both views.

on both new add-on hardware and software. To overcome these time consuming developments we therefore opted to expand the verified design of the core hardware board to a new design that is capable of connecting five three element photodiode add-on boards with an RJ12 connector. This new design has the same microcontroller at its core but differs to the original design in a few ways. Since there would be multiple add-on photodiode boards connected to it by a RJ12-cable, which are each considered to be individual mobile nodes, the 9-DoF IMU sensor data would not produce useful information and was therefore removed from this design. With the rise of Internet-of-Things (IoT) there is an abundance of wireless connectivity System-on-Chips (SoCs) and modules, the added ATWINC1500 module is one of these IoT devices that introduces WiFi networking capabilities and can provide our second measurement system implementation an untethered data connection to a multitude of network attached media instead of using the USB-connection to a computer. When connected the USB connection both powers the board and charges a single cell Lithium-Ion or Lithium-

Polymer battery that provides the system of power when used wirelessly. Figure 8 shows the schematic overview of the expanded design of the core hardware board that can connect to five add-on photodiode boards. The add-on photodiode

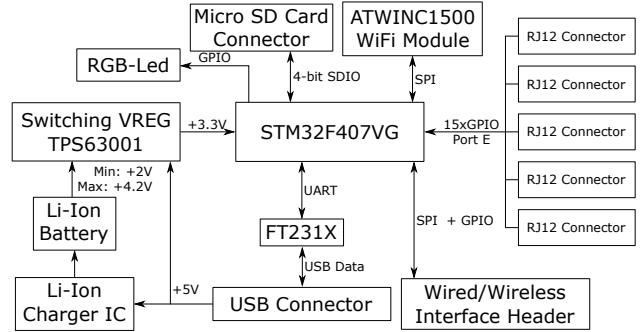


Fig. 8. Schematic overview of the core expanded hardware board. This design is capable of connecting to maximum of five photodiode boards which in this system are considered to be the mobile receiver nodes.

boards for the expanded design have an identical schematic to the boards that can be used in the hardware stack except for the RJ12-connectors for interfacing the outputs of the two-stage amplifiers to the microcontroller input pins. This type of connector was chosen for its convenient and reliable way of connecting the boards. The assembled boards are shown in Figure 9.

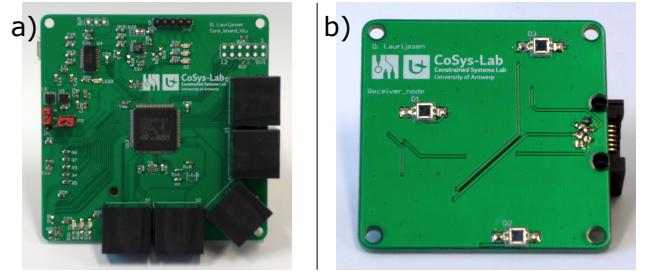


Fig. 9. a) The developed expanded core hardware board (7.5cm x 7.5cm) with the RJ12-connectors for connecting up to five photodiode boards. b) A developed photodiode board (5cm x 5cm) fitted with an RJ12-connector for connecting to the expanded core hardware board.

#### IV. PROBABILISTIC POSE ESTIMATION

When performing the actual pose estimations, one of the inherent conditions for this type of pose estimation systems is line-of-sight between the emitter and receiver. Since occlusions could occur when using this system for human body pose estimations, the problem of possible occlusions has to be dealt with in combination with inherent inaccuracies in the sensor measurements. Therefore we propose using a probabilistic approach for estimating the pose of each individual mobile receiver node. This approach has the advantage of being more robust to missed or incomplete measurements with respect to a deterministic method, which allows the use of recursive Bayesian filtering techniques such as extended Kalman filters or particle filters. Since these techniques can

cope with incomplete data sets for updating the pose estimate of each node, as long as the required minimum of independent measurements is obtained. This presents the opportunity for creating a probabilistic framework in which multiple different sensors can co-operate. Every sensor will present its sensor data to the estimator which in turn will process the data and update or refine the estimated pose independent from the other sensors. An example of such a scenario would be using the on-board 9-DoF IMU-sensor of the hardware stack in combination with the infrared sensor where the IMU can refine the estimated orientation of the node or update the entire pose in case of occlusions.

Since we will only work with the received infrared Lighthouse emissions in this paper, we will focus on the foundation of the probabilistic framework that currently processes this type of data for performing the actual pose estimations. When processing the discretized received signals from the nodes every photodiode ( $PD_i$ ,  $i = 1, 2, 3$ ) is positioned in 3D-space by an angle in the horizontal plane (azimuth angle  $\phi$ ) and an angle in the vertical plane (elevation angle  $\theta$ ) with regard to the position of the Lighthouse ( $LH_A$  or  $LH_B$ ) that emitted the infrared laser sweep. This returns a measurement  $\mathbf{M}$  that comprises the individual azimuth and elevation measurements obtained for each lighthouse-photodiode pair:

$$\mathbf{M} = \begin{bmatrix} \phi_{LH_A-PD_1} & \phi_{LH_A-PD_2} & \phi_{LH_A-PD_3} \\ \theta_{LH_A-PD_1} & \theta_{LH_A-PD_2} & \theta_{LH_A-PD_3} \\ \phi_{LH_B-PD_1} & \phi_{LH_B-PD_2} & \phi_{LH_B-PD_3} \\ \theta_{LH_B-PD_1} & \theta_{LH_B-PD_2} & \theta_{LH_B-PD_3} \end{bmatrix}$$

Since every mobile receiver node has three photodiodes that are fixed on the PCB, of which we know the exact relative positions, the photodiode positions will always be in the same plane regardless of its orientation. The array configuration in combination with the relative positions of the Lighthouses is crucial a priori information for the proper operation of the probabilistic framework for estimating the full pose of the mobile nodes.

For a certain pose  $\vec{P}$  that defines a position ( $X, Y, Z$ ) and orientation ( $\alpha, \beta, \gamma$ ) of the photodiode array can, in combination with the geometry of the array, be used to calculate the azimuth and elevation angles that each diode would be positioned at with respect to both lighthouses. This results in a calculation matrix  $\mathbf{C}(\vec{P})$  which is a function of the array pose  $\vec{P}$ , and comprises, similar to the measurement matrix  $\mathbf{M}$ , the azimuth and elevation angles for every photodiode-lighthouse pair:

$$\mathbf{C}(\vec{P}) = \begin{bmatrix} \phi_{LH_A-PD_1} & \phi_{LH_A-PD_2} & \phi_{LH_A-PD_3} \\ \theta_{LH_A-PD_1} & \theta_{LH_A-PD_2} & \theta_{LH_A-PD_3} \\ \phi_{LH_B-PD_1} & \phi_{LH_B-PD_2} & \phi_{LH_B-PD_3} \\ \theta_{LH_B-PD_1} & \theta_{LH_B-PD_2} & \theta_{LH_B-PD_3} \end{bmatrix}$$

Assuming Gaussian error distributions on the azimuth and elevation measurements we can define a likelihood function

$\mathcal{L}(M|\vec{P})$  for a measurement  $\mathbf{M}$  given a pose  $\vec{P}$  as:

$$\mathcal{L}(M|\vec{P}) = \exp\left(-\frac{1}{2} \cdot (\vec{M}_v - \vec{C}_v)^T \cdot \Sigma^{-1} \cdot (\vec{M}_v - \vec{C}_v)\right) \quad (1)$$

with  $\vec{M}_v$  and  $\vec{C}_v$  being vectorized versions of the matrices  $\mathbf{M}$ ,  $\mathbf{C}(\vec{P})$  and  $\Sigma$  being the distribution's covariance matrix. We can define a posterior probability function for the pose  $\vec{P}$  given measurement  $\mathbf{M}$  through Bayes rule:

$$P(\vec{P}|\mathbf{M}) = \frac{\mathcal{L}(\mathbf{M}|\vec{P}) \cdot P(\vec{P})}{P(\mathbf{M})} \quad (2)$$

with  $P(\vec{P})$  the prior distribution for pose  $\vec{P}$  and  $P(\mathbf{M})$  the marginal distribution of the product in the numerator. The prior  $P(\vec{P})$  allows constraining the solution space to sensible poses (e.g. the array will always be oriented in an upright orientation). In this paper we will minimize the negative logarithm of the posterior distribution to arrive at a pose estimate for the sensor:

$$\vec{P}_{est} = \arg \min(-\log(P(\vec{P}|\mathbf{M})) \quad (3)$$

The posterior is minimized using an unconstrained non-linear minimization function (Nelder-Mead). The covariance matrix is set to be a scaled identity matrix:

$$\Sigma = \sigma \cdot I \quad (4)$$

with  $\sigma$ , which represents the measurement noise on the azimuth/elevation estimates, set to 2.5 degrees.

## V. EXPERIMENTAL RESULTS

In order to verify whether the proposed hardware and probabilistic framework are capable of estimating the full pose of the mobile receiver node(s) using the HTC Lighthouses, two different experiments were performed. In the first experiment only one mobile receiver node was used to perform the measurements. The main goal of this experiment is to find the accuracy and precision of full pose estimations. The second experiment uses multiple receiver nodes simultaneously that are strapped to a human test subject. In a first attempt the goal is rather to estimate a partial body pose using the proposed multi-node system to illustrate the applicability of the proposed approach for human pose estimation.

### A. Single-Node Human Body Pose Measurements

For performing the single node pose estimations a measurement jig was created that fixed the Lighthouses on their given positions in the world coordinate system. The mobile receiver node, for which the hardware stack measurement system was used, can be positioned on the jig at intervals of 5cm in  $X$ , at intervals of 10cm in  $Y$  and could be rotated about the  $Z$ -axis using a round protractor to adjust the  $\gamma$  of the pose. The position of the  $Z$ -axis was otherwise fixed together with the  $\alpha$  and  $\beta$  orientations. Although the measurement jig itself has a slight deviation due to fabrication tolerances, it allows for reliably testing the efficacy of the probabilistic framework. For demonstrating the accuracy and precision of this single node pose estimation system three different poses

(a, b and c) were set up in the measurement jig for which 200 measurement were performed per pose. In Figure 10 the Lighthouses are represented by two squares at their respective position and orientation. The reference and estimated position for each node have been drawn as the center of the axes with their orientation represented by the direction of the axes. The reference pose is drawn with full lines and the estimated pose is drawn with dashed lines. Since a 3D-situation is difficult to interpret in a 2D-representation, the difference between the true and estimated positions are represented with the projected full lines on the  $XY$ -plane. Figure 10a), 10b) and 10c) show a close-up of each pose in a similar fashion. For expressing the accuracy of the pose estimation system, the Mean Absolute Error (MAE) has been chosen. These MAE values are shown in Table I for every pose. Similarly the standard deviation has been used for expressing the precision of the pose estimation system, which can be found in Table II. The values in these tables demonstrate the efficacy of the proposed system since the maximum MAE for the positions is 31.2mm with a maximum standard deviation of 0.4mm and the maximum MAE for the orientations is  $5.44^\circ$  with a maximum standard deviation of  $0.88^\circ$ .

TABLE I  
MEAN ABSOLUTE ERROR FOR SINGLE NODE POSE ESTIMATION

	MAE Pose It. a	MAE Pose It. b	MAE Pose It. c
$X$	18.9mm	7.4mm	6.9mm
$Y$	15.8mm	26.1mm	31.2mm
$Z$	17.9mm	18.4mm	16.2mm
$\alpha$	$0.49^\circ$	$0.91^\circ$	$0.17^\circ$
$\beta$	$1.85^\circ$	$2.18^\circ$	$0.74^\circ$
$\gamma$	$5.44^\circ$	$3.31^\circ$	$0.88^\circ$

TABLE II  
STANDARD DEVIATIONS FOR SINGLE NODE POSE ESTIMATION

	Standard Dev. Pose a	Standard Dev. Pose b	Standard Dev. Pose c
$X$	0.3mm	0.3mm	0.4mm
$Y$	0.1mm	0.2mm	0.2mm
$Z$	0.1mm	0.1mm	0.2mm
$\alpha$	$0.15^\circ$	$0.18^\circ$	$0.26^\circ$
$\beta$	$0.35^\circ$	$0.46^\circ$	$0.74^\circ$
$\gamma$	$0.39^\circ$	$0.88^\circ$	$0.70^\circ$

### B. Multi-Node Pose Measurements

Given that the goal for the proposed measurement system is to estimate a 6-DoF pose for multiple nodes in order to establish a human body pose estimation system a second experiment was performed to illustrate the applicability of the proposed approach. In this experiment the expanded core hardware board was used in combination with the five photodiode add-on boards each connected by an RJ12-cable. The core

hardware board and every photodiode board were strapped to the plexus and both the lower and upper legs of a human test subject. For a first iteration of this experiment the test subject stood in the measurement environment, equipped with the two Lighthouses fixed at known positions and orientations, in a resting pose (as shown on the right of Figure 11). In this pose 30 measurements were performed for every mobile receiver node simultaneously and processed using the probabilistic framework. On the left of Figure 11 the estimated poses of all the mobile nodes are shown as green cuboids, these poses are used to generate the lower limbs and the orientation of the upper body of a stick figure. When the left and right image of Figure 11 are compared, these two appear to closely resemble each other. The variations of the pose estimates in the multi-node system were not similar to the measured pose variations with the single-node system from the previous section. For a second iteration of this experiment the test subject had to raise its right leg with the same measurement system configuration strapped to its body (as shown on the right of Figure 12). For this iteration 30 measurements were performed as well which return the mean estimated poses of the mobile nodes shown on the left of Figure 12 once processed. These mean estimated poses were used to calculate and draw the stick figure's lower limbs and orientation of the upper body in Figure 12, it can be seen that this closely resembles the real pose of the subject as well.

### VI. DISCUSSION, CONCLUSIONS AND FUTURE WORK

When combining the measurement results of the mean absolute errors and the standard deviations of the single-node pose estimations together with the measurement results of the multi-node pose estimations, it can be stated that the proposed measurement system already shows promising results towards establishing an affordable yet precise and accurate six degrees-of-freedom human body pose estimation system. Improvements however are still in need to make this system more reliable. Although using the proposed probabilistic framework already provides a more robust solution, one of the major concerns using this type of measurement system is a full occlusion (e.g. when the node would be perpendicular to the front of the Lighthouses). This problem could be partially alleviated by adding the 9-DoF IMU sensor data (already integrated on the core hardware board of the hardware stack) to the probabilistic framework. This solution however only works for a limited amount of time since IMU-sensors suffer from an inherent integration drift. The other approach would be to use more Lighthouses in the environment to prevent occlusions. However this also raises the system cost.

Given our previous work in the field of ultrasonics another addition to our proposed measurement system would be ultrasonic Time of Flight measurements. These measurements would return the distance between an ultrasound emitting source (at a given reference point) and a microphone (array) that would be added to the mobile receiver node(s). These measurements would elegantly integrate into the proposed probabilistic framework and enable us to define the estimated

## Single Node Pose Measurements

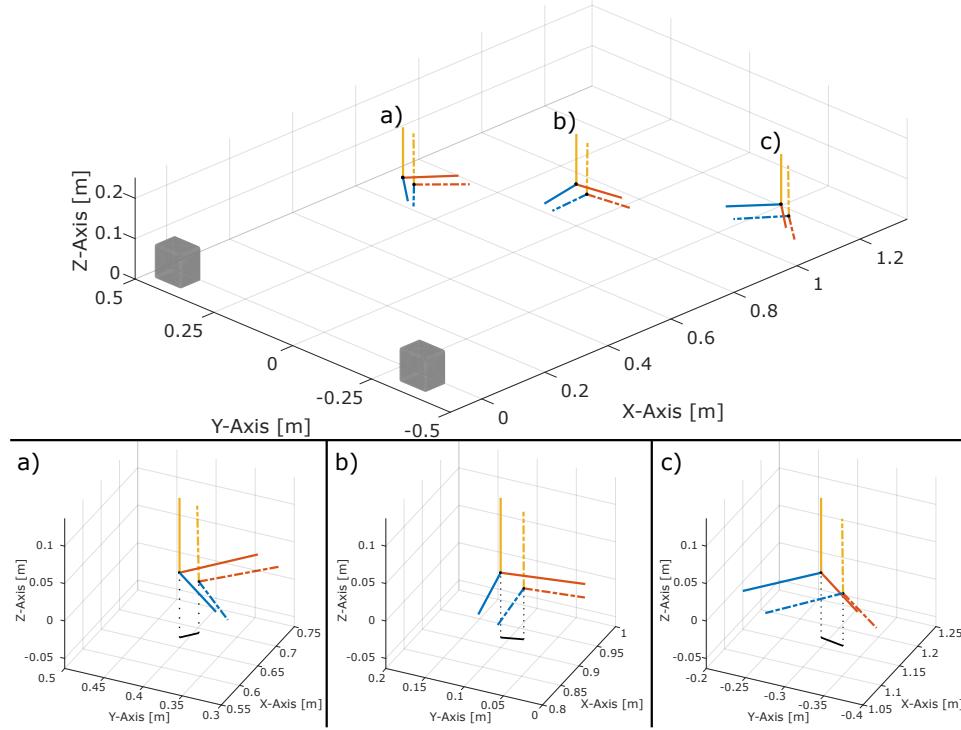


Fig. 10. At the top of this figure the two Lighthouses, used to perform these measurements, are shown as two cubes at their respective fixed world-coordinates and given orientation. A hardware stack measurement system was used as the mobile receiver node at three different positions and orientations (a, b and c). The reference and estimated position of every iteration of the mobile receiver node are shown as the center of the plotted axes where the orientation is represented by the direction of the axes. The full lines of these axes represent the reference pose where the axes with the dotted lines represent the estimated pose. In order to better visualize the difference between the reference and estimated position for every pose the connection line between these positions is projected on the  $XY$ -plane. A close-up of every pose is shown in a), b) and c).

position of mobile nodes in a spherical coordinate system (azimuth, elevation and distance).

In addition to the IMU-sensor and ultrasonic data integration to the probabilistic framework we wish to work towards a framework that will constrain the estimated poses by the limits of the human body. In the current human body pose estimation system a very simple stick-based human body model was used, which only suffices for illustrative purposes. Since this model is unrealistic, it will be replaced by more advanced biomechanical models such as the Biomechanics of Bodies (BoB model [19], [20]. In addition to more advanced human body models the probabilistic framework will be redefined in such way that the independent variables of the posterior represent the degrees of freedom of the human body model instead of the individual sensor poses. The updated posterior will allow for directly optimizing the body pose instead of fitting a body pose to the individually estimated sensor poses. This approach should produce more reliable and robust body pose estimations and allow for efficient fusion with other sensors such as sonar and IMU.

### ACKNOWLEDGMENTS

This work was in part enabled by a Special Research Grant (BOF-STIMPRO) of the University of Antwerp.

### REFERENCES

- [1] J. Egger, M. Gall, J. Wallner, P. de Almeida Germano Boechat, A. Hann, X. Li, X. Chen, and D. Schmalstieg, "Integration of the HTC Vive into the medical platform MeVisLab," T. S. Cook and J. Zhang, Eds. International Society for Optics and Photonics, mar 2017, p. 1013817.
- [2] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments," *The International Journal of Robotics Research*, vol. 31, no. 5, pp. 647–663, apr 2012. [Online]. Available: <http://ijr.sagepub.com/cgi/doi/10.1177/0278364911434148>
- [3] K. Khoshelham and S. O. Elberink, "Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications," *Sensors*, vol. 12, no. 12, pp. 1437–1454, feb 2012. [Online]. Available: <http://www.mdpi.com/1424-8220/12/2/1437>
- [4] I. Oikonomidis, N. Kyriazis, and A. A. Argyros, "Efficient Model-based 3D Tracking of Hand Articulations using Kinect." [Online]. Available: <http://www.ics.forth.gr/>
- [5] Y.-J. Chang, S.-F. Chen, and J.-D. Huang, "A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities," *Research in Developmental Disabilities*, vol. 32, no. 6, pp. 2566–2570, 2011.
- [6] J. Preis, M. Kessel, M. Werner, and C. Linnhoff-Popien, "Gait Recognition with Kinect."
- [7] M. Gabel, R. Gilad-Bachrach, E. Renshaw, and A. Schuster, "Full body gait analysis with Kinect," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, aug 2012, pp. 1964–1967. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6346340>
- [8] R. A. Clark, Y.-H. Pua, A. L. Bryant, and M. A. Hunt, "Validity of the Microsoft Kinect for providing lateral trunk lean feedback during gait retraining," 2013.

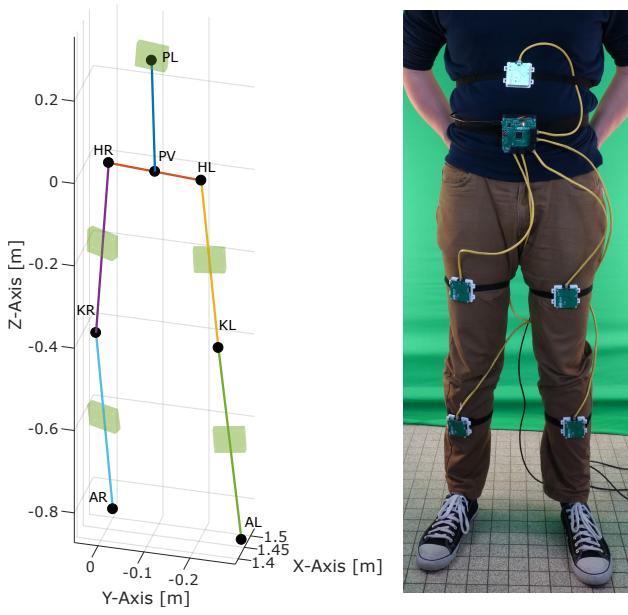


Fig. 11. Results of a human body pose estimation experiment where multiple mobile receiver nodes were strapped to the pectoral and both the upper and lower legs of a test subject in rest in order to estimate its respective poses. On the left of this figure the mean estimated poses for every node are shown as green cuboids, using these poses a stick figure is drawn that closely resembles the actual pose of our test subject, shown on the right picture.

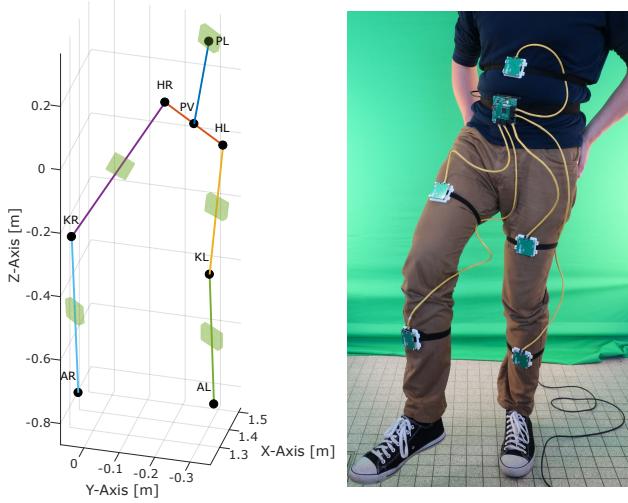


Fig. 12. Results of a second iteration of this experiment. The test subject raised its right leg (as shown above on the right picture) with the same configuration of mobile receiver nodes that are strapped to its body. When looking at the node position and orientations in combination with the generated stick figure this very much resembles the pose of our test subject.

- [9] A. Staranowicz, G. R. Brown, and G.-L. Mariottini, "Evaluating the accuracy of a mobile Kinect-based gait-monitoring system for fall prediction," in *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments - PETRA '13*. New York, New York, USA: ACM Press, 2013, pp. 1–4. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2504335.2504396>
- [10] F. Ahmed, P. P. Paul, and M. L. Gavrilova, "DTW-based kernel and rank-level fusion for 3D gait recognition using Kinect," *The Visual Computer*, vol. 31, no. 6-8, pp. 915–924, jun 2015. [Online]. Available: <http://link.springer.com/10.1007/s00371-015-1092-0>
- [11] A. Pfister, A. M. West, S. Bronner, and J. A. Noah, "Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis," *Journal of Medical Engineering & Technology*, vol. 38, no. 5, pp. 274–280, jul 2014. [Online]. Available: <http://www.tandfonline.com/doi/full/10.3109/03091902.2014.909540>
- [12] D. Vlasic, R. Adelsberger, G. Vannucci, J. Barnwell, M. Gross, W. Matusik, and J. Popović, "Practical motion capture in everyday surroundings," *ACM Transactions on Graphics*, vol. 26, no. 3, p. 35, jul 2007. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1276377.1276421>
- [13] E. Mirek, M. Rudzińska, and A. Szczudlik, "The assessment of gait disorders in patients with Parkinson's disease using the three-dimensional motion analysis system Vicon," *Neurologia i neurochirurgia polska*, vol. 41, no. 2, pp. 128–33, jan 2007. [Online]. Available: <http://europepmc.org/abstract/med/17530574>
- [14] M. Cognolato, "Experimental validation of Xsens inertial sensors during clinical and sport motion capture applications," Ph.D. dissertation, University of Padova, Padova, Italy, 2012. [Online]. Available: [http://tesi.cab.unipd.it/40443/1/Tesi\\_Completa.pdf](http://tesi.cab.unipd.it/40443/1/Tesi_Completa.pdf)
- [15] T. Seel, J. Raisch, and T. Schauer, "IMU-Based Joint Angle Measurement for Gait Analysis," *Sensors*, vol. 14, no. 4, pp. 6891–6909, apr 2014. [Online]. Available: <http://www.mdpi.com/1424-8220/14/4/6891>
- [16] D. Laurijssen, S. Truijen, W. Saeys, W. Daems, and J. Steckel, "A flexible embedded hardware platform supporting low-cost human pose estimation," in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, oct 2016, pp. 1–8. [Online]. Available: <http://ieeexplore.ieee.org/document/7743682/>
- [17] ———, "An Ultrasonic Six Degrees-of-Freedom Pose Estimation Sensor," *IEEE Sensors Journal*, pp. 1–1, 2016. [Online]. Available: <http://ieeexplore.ieee.org/document/7592877/>
- [18] D. Laurijssen, S. Truijen, W. Saeys, and J. Steckel, "Three sources, three receivers, six degrees of freedom: An ultrasonic sensor for pose estimation & motion capture," in *2015 IEEE SENSORS*. IEEE, nov 2015, pp. 1–4.
- [19] B. May, J. Shippen, and A. Woodcock, "Contemporary Ergonomics and Human Factors 2013: Proceedings of the international conference on Ergonomics & Human Factors 2013," 2013. [Online]. Available: <https://pureportal.coventry.ac.uk/en/publications/biomechanical-analysis-of-the-walking-of-encumbered-and-unencumbe-2>
- [20] J. Shippen and B. May, "Biomechanical modelling system to solve muscular loading distribution," 2008. [Online]. Available: <https://pureportal.coventry.ac.uk/en/publications/biomechanical-modelling-system-to-solve-muscular-loading-distribu-2>