

**This item is the archived peer-reviewed author-version of:**

A flexible embedded hardware platform supporting low-cost human pose estimation

**Reference:**

Laurijssen Dennis, Truijen Steven, Saeys Wim, Daems Walter, Steckel Jan.- A flexible embedded hardware platform supporting low-cost human pose estimation

2016 International Conference on Indoor Positioning and In door Navigation (IPIN), 4-7 October, 2016, Madrid, Spain - ISSN 2162-7347 - IEEE, 2016, p. 1-8

Full text (Publishers DOI): <http://dx.doi.org/doi:10.1109/IPIN.2016.7743682>

To cite this reference: <http://hdl.handle.net/10067/1372950151162165141>

# A Flexible Embedded Hardware Platform Supporting Low-Cost Human Pose Estimation

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

**Abstract**—Throughout the last decades, human motion capture systems have become an important tool for various sectors. Besides the entertainment sector, which is probably the most known sector to use these systems through popular movies and video games, the medical sector has adopted this technology as an analysis tool. These clinical analyses measure the human body posture and movement for various purposes including rehabilitation and sports, and require high accuracy and precision. Although there is a variety of products that offer these capabilities, there is not yet an affordable system that meets the requirements for medical implementations. We therefore propose using an ultrasonic six degrees-of-freedom sensor system which combines a distributed array of ultrasonic emitters and a small array of ultrasonic receivers. In order to support low-cost human pose estimation, while retaining the possibility for re-use and expandability, the need for a flexible embedded hardware platform naturally emerged. The proposed hardware design supports the requirements of a high accuracy and high precision pose estimation system while keeping its overall system cost low.

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

## I. INTRODUCTION

Human motion monitoring has taken a great leap over the last few decades and is becoming a more common tool in a variety of sectors. The most popular sector that uses human motion monitoring, pose estimation or motion capture systems is the entertainment sector. Movies and video games have adopted and drive the innovation for these systems to further improve the accuracy and precision for making their products display ever more realistic and lifelike visual computer generated images. Although these systems provide the entertainment studios with highly accurate and precise pose estimations, these systems have a high system cost and usually rely on a dedicated measurement environment. With the introduction of the Nintendo Wii video game console, video games have evolved to use basic human motion as an input device besides standard action or directional buttons and joystick movements. Other companies followed in Nintendo's footsteps by making the PlayStation Eye and the Microsoft Kinect available for their respective consoles. The Kinect provided an affordable tool for several research topics such as (3D) indoor mapping [1], [2], detection of humans [3],

human pose estimation [4], [5] and gait analysis [6]–[8]. These last two demonstrate that the medical sector has also adopted human motion monitoring as a tool that serves purposes such as rehabilitation and sports performances.

For the medical sector, body pose estimation used for gait analysis require high accuracy and high precision measurements thus relying on expensive measurement systems (e.g. ViCon or Xsens [8]–[11]). Since affordable products (e.g. the Microsoft Kinect) are not yet up to the task of providing clinical implementations [12], movement evaluations are of limited availability and remain costly to patients suffering from a movement dysfunction. The goal of our research is to develop an affordable yet accurate and precise human body pose estimation system that meets the requirements for movement analysis. In order to meet the aforementioned requirements, we previously proposed to use an ultrasonic sensor array [13] to develop a low-cost six Degrees-of-Freedom (DOF) sensor system that is able to estimate the position ( $X, Y, Z$ ) of a sensor in a Cartesian world coordinate system and the rotation around the three principle axes ( $\alpha, \beta$  and  $\gamma$ ).

Because of promising results of this human motion monitoring system, which have been demonstrated in previous work through simulation and a proof of concept prototype using both off the shelf and available hardware [13], the necessity for an improved custom hardware platform has emerged. Since this hardware platform is intended for human body pose estimation measurements, the hardware needs to be designed while keeping specific technological requirements (size, power, weight) in mind. Furthermore, it also needs to serve as a research tool for pose estimation applications thus creating the need for flexibility.

In the remainder of this paper, the overall system topology is presented, followed by a section with a thorough description of the hardware architecture that has been used for developing the proposed flexible hardware platform. In the subsequent section, the prototype of the hardware platform is presented with its extensions. The results of an experimental setup using the latter is given in the following section. In the final section, we will discuss our proposed system, present possible extensions and hardware revisions as future work together with the conclusions of this paper.

## II. SYSTEM TOPOLOGY

To perform human body pose estimations, the ultrasonic sensor system relies on Time-of-Flight (ToF) Measurements where a transducer subsystem emits the ultrasonic pulses and a receiver subsystem receives the broadcast pulses. This receiver subsystem will be referred to as the *mobile (receiver) node* in the remainder of this paper. In order to estimate the pose of the mobile receiver node, at least three transducers are required in combination with at least three microphones on the mobile receiver node [14]. Figure 1 shows a schematic overview of the proposed system topology using three transducers  $T_1, T_2$  and  $T_3$  which broadcast ultrasonic pulses simultaneously. In order to determine the time these pulses travel from each transducer to the microphones,  $M_1, M_2$  and  $M_3$  (fitted onto the mobile receiver node), matched filtering techniques are used. The efficacy of this technique relies on the auto and cross correlation properties of the emitted signals. Due to the properties of a time series of independent identically distributed random variables (shown in Figure 2), also known as white noise, pseudo-random signals are very suited once they have been band limited to fall within the spectral passband range of the transducers. Every microphone signal is fed into the matched filter bank, which contains the reference signals of the three transducers, thus yielding three peaks in the matched filter bank outputs which represent the arrival times or the times of flight. The operation of this technique is demonstrated in Figure 3. Using the output of the matched filter bank in combination with the speed of sound (which propagates through air at an approximate speed of 340m/s), the distance between every microphone-transducer pair can be calculated. These distances form the distance vector that is used for estimating the pose of the mobile receiver node in the environment using a maximum likelihood estimator in combination with a multidimensional constrained nonlinear minimization function [13].

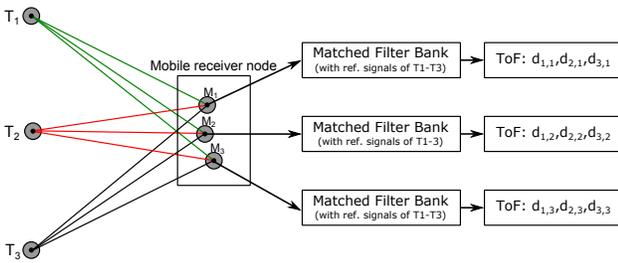


Fig. 1. A schematic overview of the system topology where  $T_1, T_2$  and  $T_3$  represent three transducers positioned in the environment and a mobile receiver node equipped with three microphones ( $M_1, M_2$  and  $M_3$ ). Once the ultrasonic pulses have been received by the microphones the signals are transferred to a matched filter bank. Every output of the matched filter banks will exhibit a peak that indicates the arrival time for the signal emitted by a transducer. These arrival times represent the ToF which is used to calculate the distance for every microphone-transducer pair.

The transducer subsystem consists of a custom embedded hardware system, custom amplifier hardware and the actual transducers. The embedded hardware incorporates a microcontroller which connects to a mobile receiver node through either

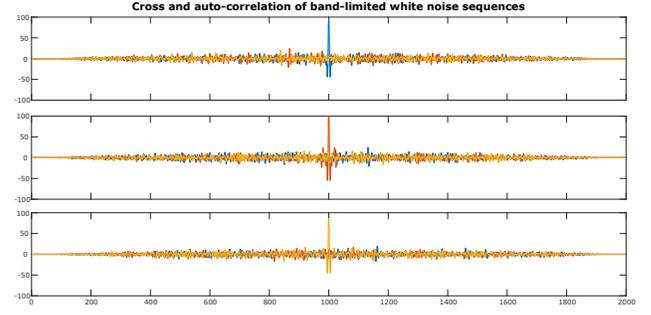


Fig. 2. The cross and auto correlation functions of three band-limited white noise sequences where a) shows the auto correlation of signal 1 (blue) and the cross correlation of signals 1 & 2 (red) and signals 1 & 3 (yellow). b) shows the auto correlation of signal 2 (red) and the cross correlation of signals 2 & 1 (blue) and signals 2 & 3 (yellow). c) shows the auto correlation of signal 3 (yellow) and the cross correlation of signals 3 & 1 (blue) and signals 3 & 2 (yellow).

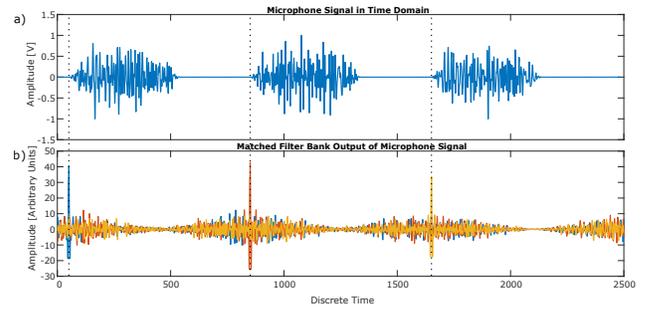


Fig. 3. a) Shows a simulated microphone signal that exhibits three white noise sequences originating from three different transducers. b) Represents the output of an FFT-based matched filter bank in which three peaks can be perceived which represent the time of arrival. These peaks correspond to the beginning of every sequence in the microphone signal. These received signal sequences have been elaborately spread in time in this picture to demonstrate the operation of the matched filter technique.

a wired or wireless connection for synchronization and Digital-to-Analog-Converters (DACs) that can each generate a different yet synchronous analog signal. The outputs from these DACs are connected to the custom amplifier hardware which amplifies the signal to the adequate level for the ultrasonic transducers (e.g. a Senscomp 7000 Ultrasonic transducer [15]) to emit the ultrasonic pulses. These transducers are placed at fixed positions in the environment forming a distributed transducer array.

The receiver subsystem or mobile receiver node is an embedded hardware system that consists of a microcontroller, a wired or wireless synchronization connectivity option to the transducer subsystem and a microphone array (of at least three elements). The design and architecture of the receiver subsystem in order to create a flexible hardware platform for validation of our proposed ultrasonic six DoF sensor system while retaining enough expansion capabilities will be elaborated in the remainder of this paper.

## III. HARDWARE ARCHITECTURE

The receiver subsystem or mobile receiver node is the element of the pose estimation sensor system that can be com-

pared to an active marker in other human motion monitoring or motion capture systems. In order to use our proposed sensor system for human body pose estimation, these mobile receiver nodes need to be attached to various body parts for which the pose will be estimated. The size and weight for the receiver nodes is therefore limited. Since the features for the mobile receiver node should allow for easy adaptation in terms of the number of microphone element within the array or a specific wireless communication protocol, implementing these features on one Printed Circuit Board (PCB) would not result in with a flexible and expandable platform. Therefore, the receiver subsystem has been designed with a core hardware board that features the microcontroller and peripherals that don't require adaptation and connects the aforementioned SPI and ADC interfaces to standard interconnection headers to ensure expandability to add-on boards. These add-on boards can be connected to either the front or the back of the core hardware board depending on the functionality (e.g. the microphone is an add-on board that connects to the front of the core hardware PCB).

#### A. Core Hardware Platform

The mobile nodes have to be constrained in size and weight yet have to be powerful enough in order to be able to perform their tasks. We have therefore chosen to use an embedded system with an ARM Cortex M4, more specifically an STM32F407VG [16] from ST Microelectronics, at its core. This 32-bit microcontroller offers high performances due to its ability to run its system core clock at 168MHz, an integrated floating point unit and the capability of executing DSP instructions. Besides these features, a large number of communication peripherals have been integrated into this microcontroller such as: two USB On-The-Go (with High-Speed support) interfaces, two dedicated I<sup>2</sup>S audio interfaces, Ethernet MAC10/100, six U(S)ART interfaces, three SPI interfaces, three I<sup>2</sup>C interfaces, two CAN-bus interfaces and an SDIO interface. Additionally, two 12-bit DACs and three 12-bit Analog-Digital Converters (ADCs), which are capable of reaching 2.4MSPS synchronously, have been integrated as well. All of these features combined with 17 timers, low-power modes, DMA and flexible static memory controllers, a true random number generator and an integrated crypto/hash processor are available in several small packages that can be as small as 4.2x4.0mm while having a reasonable cost.

Due to the great amount of features, it offers great flexibility for our mobile receiver node to meet the requirements of our hardware platform, i.e. the capability of connecting to a small microphone array and interfacing the receiver subsystem to the transducer subsystem through either a wired or wireless connection. In order to meet the requirements of the former, the three available ADC interfaces are used allowing the system to sample the microphone samples simultaneously. The wireless connectivity between the transducer subsystem is established using the SPI interface, which is a common communication protocol for a great variety of wireless transducer Integrated Circuits (ICs) or even wireless System-on-Chips (SoCs) that

could be used as a co-processor. Basic GPIO pins provide for an easy and fail-safe method for a wired connection to the transducer subsystem.

The architecture of the core hardware board is presented in a schematic overview in Figure 4. For interfacing with a computer and powering the mobile node, a nowadays very common USB connector has been added. Although the STM32F407 has two USB OTG interfaces, we have chosen to interface the USB data connection to an FTDI FT232H Single Channel Hi-Speed USB to UART IC [17] to avoid using USB drivers on the microcontroller since these middleware drivers can interfere with the interrupt priorities programmed in the measurement routines and with microcontroller debugging. This Hi-Speed USB to UART IC forms a bridge between the USB data and the microcontroller through the fastest available UART interface. Using these components we are able to set up a serial connection between the STM32F4 and a computer with a maximum baud rate of 10.5Mbaud. Using this connection the hardware platform can rapidly and easily transfer (sensor) data over to a computer where it can be further processed in applications such as Matlab.

By combining the ultrasonic distance measurements with the measurements of an Inertial Measurement Unit (IMU) sensor using a sensor fusion algorithm, the update rate, precision and accuracy of the pose estimations will increase. Therefore an LSM9DS0 9-DoF IMU from ST Microelectronics has been added to the core hardware board that features a 3D digital linear acceleration sensor (accelerometer), a 3D digital angular rate sensor (gyroscope) and a 3D digital magnetic sensor (magnetometer or digital compass) [18]. Adding this sensor to the hardware platform adds inherent pose estimation capabilities to the mobile receiver node since this device is capable of measuring the rate of acceleration ( $\pm 2g/\pm 4g/\pm 6g/\pm 8g/\pm 16g$ ) and rotation ( $\pm 245/\pm 500/\pm 2000$  dps) together with magnetic field measurements ( $\pm 2/\pm 4/\pm 8/\pm 12$  Gauss). The disadvantage of these IMUs, as a stand-alone sensor, however is that they suffer from inherent integration drift which is an undesirable effect when using these sensors for human motion monitoring systems. In order to cancel out these accumulated errors, calibration routines need to be applied. In this specific case, the ultrasonic measurements could serve as reference for the IMU to correct for the integration drift errors. Using this IMU in a sensor fusion framework will be addressed in the near future.

Besides using the accommodated Hi-Speed USB connection for transferring data, the integrated SDIO interface has been used to equip the hardware platform with a micro SD card connector. This option offers the ability to store sensor data on a non-volatile storage medium, which have become very cheap over the last few years, in either a raw file format or a human readable file format using an embedded file system (e.g. FatFs). Additionally this interface provides a more stand-alone capability to the hardware platform since all the data can be stored while being powered by an external power source (e.g. a battery pack). These external power sources can range from +1.8V to +5.5V and can even be applied

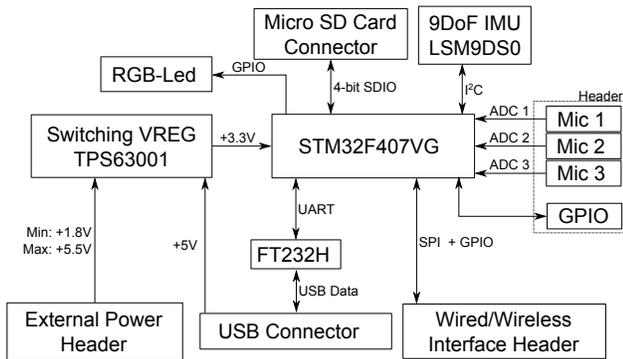


Fig. 4. Schematic of the hardware architecture of the mobile receiver node with the STM32F407VG at the core.

to the hardware platform while connected through a USB connection. In order for these sources to create a stable power supply for the microcontroller and its peripherals, a TPS63001 IC from Texas Instruments [19] is used. This high-efficiency switching Buck-Boost voltage convertor has a fixed output of +3.3V and can deliver current up to a maximum of 1.8A. Due to the relatively wide range of input voltages, this voltage convertor IC is suited to be used in combination with rechargeable single cell lithium-ion or lithium-polymer batteries (which operate between +2.25V and +4.2V) and USB power (which operates between +4.5V and +5.5V). The power consumption of the core hardware platform, which has currently been implemented without any low-power options (e.g. the STM32F4 that is clocked at 144MHz), ranges from 40mA to 140mA depending on the input voltage, which leaves plenty of overhead for powering add-on hardware boards through the voltage convertor IC.

### B. Add-on Hardware

As stated above, the SPI and ADC interfaces are connected to interconnection headers which allow the core hardware board to have add-on boards which can be connected to the front or back of the core board. For the proposed ultrasonic 6 DoF sensor system, we have developed a front add-on board which incorporates three Knowles Electronics SPU0410HR5H [20] microphones. The analog output signal of each microphone is filtered using a passive high-pass filter before being amplified using an Analog Devices AD8618 precision operational amplifier IC [21] in a non-inverting configuration. The amplified microphone signal is routed to the header connecting to the three ADC interfaces on the STM32F407. For future validation purposes, three infrared Light Emittable Diodes (LEDs) have been added to this add-on board enabling a ViCon system to perceive these LEDs as markers thus providing a means for benchmarking the proposed ultrasonic pose estimation system with a commercially available high-accuracy motion capture system.

As a back add-on board, two types of expansion boards have been developed in order to establish a wireless connection to the transducer subsystem for synchronization of the ultrasonic measurements. The first type features a Si4460 transceiver

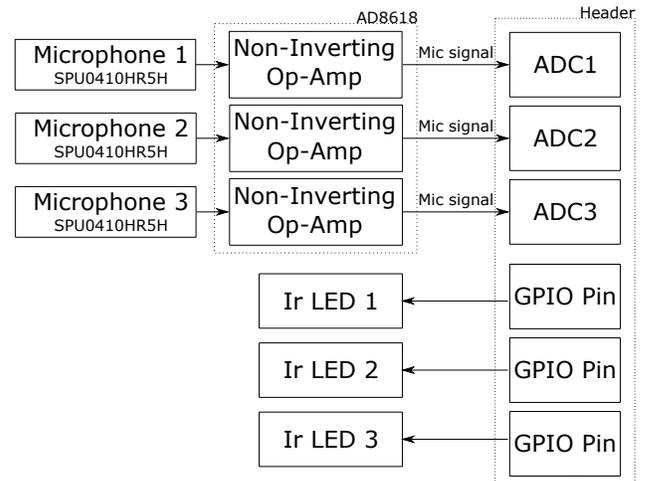


Fig. 5. Schematic of the hardware architecture of the microphone expansion board.

covering the sub-GHz frequency bands from 119–1050MHz from Silicon Labs [22] which spurs low-power usage and high sensitivity at a low cost. The transceiver provides the ability to use a range of wireless communication protocols such as the 802.15.4g WPAN network, Wireless MBus or DASH7. The second type of back add-on board incorporates a ScenSor DWM1000 module from decaWave [23] which is based on decaWave’s DW1000 Ultra Wideband (UWB) transceiver IC. The advantage of using these modules is that all of the external components for power management, clock circuitry and antenna have already been provided and that these transceivers have built-in 2-way ranging capabilities (using Time Difference of Arrival) achieving location estimates, which are claimed to have a precision up to 10 cm. This ranging feature can be integrated into the sensor fusion algorithm for even better accuracy and precision of the pose estimates in combination with the IMU and ultrasonic measurement. The disadvantage of these modules however is a relatively high cost. Both types also feature a stand-alone linear Li-Ion or Li-Polymer charge management controller IC (MCP73833 from Microchip [24]) which uses the USB power source to safely charge either the on-board lithium battery coin cell (LIR2450 [25]) or an external single cell lithium battery. If provided, either one of these batteries can power the receiver subsystem through the TPS63001.

## IV. EXPERIMENTAL PROTOTYPE

The aforementioned core hardware board of the system topology has been designed to fit a printed circuit board measuring 5cm by 5cm which is pictured in Figure 6. After a verified prototype was assembled, the design was produced in a small batch of 20 boards. In this small batch the total cost per core hardware board is €73.89 which is well within the range for low-cost devices. The price per board will drop when ordering in larger quantities.

In order to perform pose estimation experiments with the core hardware boards, the front add-on microphone board



Fig. 6. Picture of the developed core hardware board (5cm by 5cm) with the (stackable) headers.

with the required amplification circuitry has been assembled together with a back add-on board that only contains the coin cell battery holder and the lithium charging controller. When stacking these board, as can be seen in Figure 7, the combined device is capable of measuring the ultrasonic pulses originating from the transducers as a battery powered stand-alone system. At this time, the processing of the microphone data is still performed in Matlab which currently introduces the necessity for connecting the hardware stack to a computer by USB for transferring the data. This however can be alleviated by using the wireless connectivity for data transfers besides synchronization. By using the DSP capabilities of the microcontroller, the ToF post-processing can be done on the hardware board which reduces the data transfer size (since only the distances or arrival times have to be sent) and eliminates the need for a USB connection to a computer when performing measurements. Although the hardware stack itself can operate as a small stand-alone battery powered measuring device, using it as such for performing human body pose estimation measurements is cumbersome. Therefore a very basic fixture has been designed and created using 3D-printing technology. The fixture (pictured in Figure 7) provides four mounting points to fasten the hardware stack, an opening for easily replacing or removing the lithium coin cell, two strap holders for future human body pose measurements. For testing purposes a central hexagon nut has been added to the fixture base for mounting the hardware stack onto a standard camera tripod or in the case of our experimental setup a FLIR pan-tilt PTU-E46 system.

## V. RESULTS

For preliminary validation purposes an experimental hardware setup has been created. This setup consists of the trans-



Fig. 7. Picture of a stack of three boards consisting of the core hardware board, the front add-on board with three microphones and a back add-on board which has been only been fitted with the charging circuitry and a coin cell holder. The stack measures 5cm by 5cm by 3,25cm).

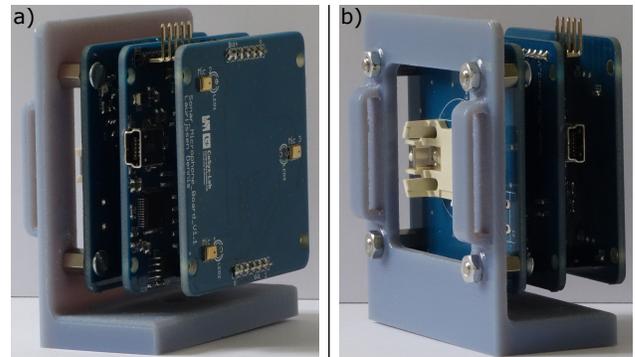


Fig. 8. a) Front view of the stack of three boards mounted on the 3D printed fixture. b) Back view of the same hardware stack mounted on the fixture.

ducer subsystem in combination with three Senscomp 7000 ultrasonic transducers and the proposed hardware platform as the mobile receiver node. The three transducers are mounted on camera tripods where the receiver node is mounted on the FLIR pan-tilt system. Since the matched filtering technique that is used for the ToF-measurements relies on reference signals, the white noise sequences generated for every transducer have to be recorded before the actual pose estimation can be performed. Therefore the signals emitted by every transducer are individually recorded by the receiver node. When the cross and auto correlation functions of the three pre-recorded transducer base signals are calculated and plotted (in Figure 9), we notice a distinct difference in the peakedness between the auto and cross correlation functions of the simulated (Fig. 2) and the recorded (Fig. 9) white noise sequences. Due to the significant non-flat spectrum of the recorded signals, caused

by non-linearities of the Senscomp transducer [15] and the transfer function of the Knowles microphones [20], the auto correlation function exhibits a lower peak-to-sidelobe ratio and a wider main lobe [26]. Furthermore the cross correlation between transducer signals retains high side lobe levels due to this non-flat signal spectrum. This impacts the accuracy of the range estimate in the case of simultaneous and overlapping transmissions and therefore our pose estimation efficacy. Once

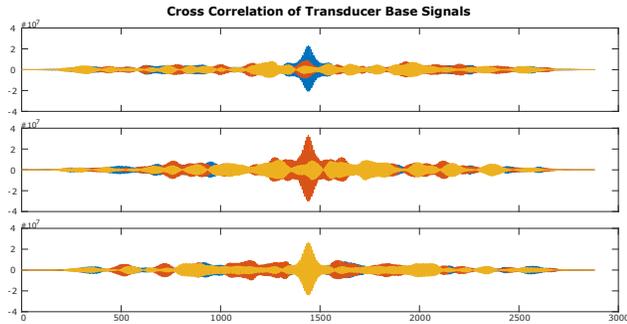


Fig. 9. The cross and auto correlation functions of the three pre-recorded transducer base signals where a) shows the auto correlation of transducer base signal 1 (blue) and the cross correlation of transducer base signals 1 & 2 (red) and transducer base signals 1 & 3 (yellow). b) shows the auto correlation of transducer base signal 2 (red) and the cross correlation of transducer base signals 2 & 1 (blue) and transducer base signals 2 & 3 (yellow). c) shows the auto correlation of transducer base signal 3 (yellow) and the cross correlation of transducer base signals 3 & 1 (blue) and transducer base signals 3 & 2 (yellow). These transducer base signals were emitted by the Senscomp 7000 Ultrasonic transducer and recorded using the Knowles microphones on the hardware platform.

all reference or *base* signals are captured, the transducers and mobile receiver node are placed in the environment. Throughout the experiment, transducers will remain in a fixed position and orientation while the mobile receiver node can vary its position and orientation by varying the azimuth and elevation of the PTU-E46 pan-tilt system. The transducer subsystem will simultaneously emit the ultrasonic pulses and trigger the mobile receiver node to start recording the incoming signals. Once these signals are quantized using the three ADC channels, the digital values are transferred to the computer for processing. The three microphone signals are extracted from the digital values which are used as the input for the matched filter bank. Using the output of the matched filtering operation in combination with the (temperature compensated) speed of sound, the ToF distances are calculated. The pose of the mobile receiver node is estimated using a maximum likelihood estimator in combination with a multidimensional constrained nonlinear minimization function. This uses the distance vector (which consists of all the ToF measurements) and specific constraints, e.g. the mobile node’s position not exceeding the dimensions of the environment.

In the experiment, the mobile receiver node assumes three poses for which it remains at the same  $XYZ$ -position but revolves around the  $Z$ -axis in steps of  $30^\circ$ . For every pose iteration, a number of measurements are performed that are used to estimate the pose of the mobile node. The results of the proposed estimation technique are shown in Figure 10 where

TABLE I  
STANDARD DEVIATIONS FOR POSE ESTIMATION ITERATIONS

	Standard Dev. Pose It. $150^\circ$	Standard Dev. Pose It. $180^\circ$	Standard Dev. Pose It. $210^\circ$
$X$	3.6mm	3.4mm	3.3mm
$Y$	3.3mm	2.8mm	10.3mm
$Z$	20.9mm	7.3mm	24.6mm
$\alpha$	$0.85^\circ$	$0.45^\circ$	$1.04^\circ$
$\beta$	$1.50^\circ$	$0.87^\circ$	$1.02^\circ$
$\gamma$	$18.70^\circ$	$11.78^\circ$	$13.67^\circ$

TABLE II  
MEAN ABSOLUTE ERROR OVER ALL POSE ESTIMATES

	Mean Absolute Error
$X$	30.4mm
$Y$	11.2mm
$Z$	70.7mm
$\alpha$	$0.99^\circ$
$\beta$	$1.43^\circ$
$\gamma$	$29.83^\circ$

three pose iterations of the true pose of the mobile receiver node have been drawn with their position as the center of the axes and their rotation represented by the direction of the axes. In a), b) and c) an ellipsoid has been drawn with the semi principal axis lengths equal to the standard deviation on the position estimate for the corresponding axis. In d), e) and f) a cone has been drawn over each axis with its opening angle equal to the standard deviation of the rotation estimate for the corresponding axis. The standard deviations for the pose iterations can be found in Table I. Table II shows the mean absolute errors for all measurements of the experiment.

## VI. DISCUSSION, CONCLUSIONS AND FUTURE WORK

When reviewing our experimental results, the proposed pose estimation system suffers from a relatively large bias in its pose estimations and a high variation in the estimation of  $\gamma$ , while exhibiting low variance in the position estimates ( $X, Y$  and  $Z$ ) and the  $\alpha, \beta$  rotation estimates. These errors are induced by the non-linear properties of the Senscomp transducers which impede the efficacy of the matched filtering technique when the transducers simultaneously emit their pulses, which in turn yields inaccurate times of flight. In order to improve these pose estimates, an improved method for determining the ToF needs to be implemented that can cope with the non-linearities of the Senscomp transducers, e.g. based on sparse representations [27]. Alternatively the transducer subsystem can be adapted to use a more linear transducer material such as Emfit [28], [29]. Using (mutual) information theory, the configuration of both the microphone and the transducer array can be optimized for their positions in the environment or the number of elements. Due to the flexibility of using the front add-on PCB, the microphone configuration can be easily tested without having to redesign

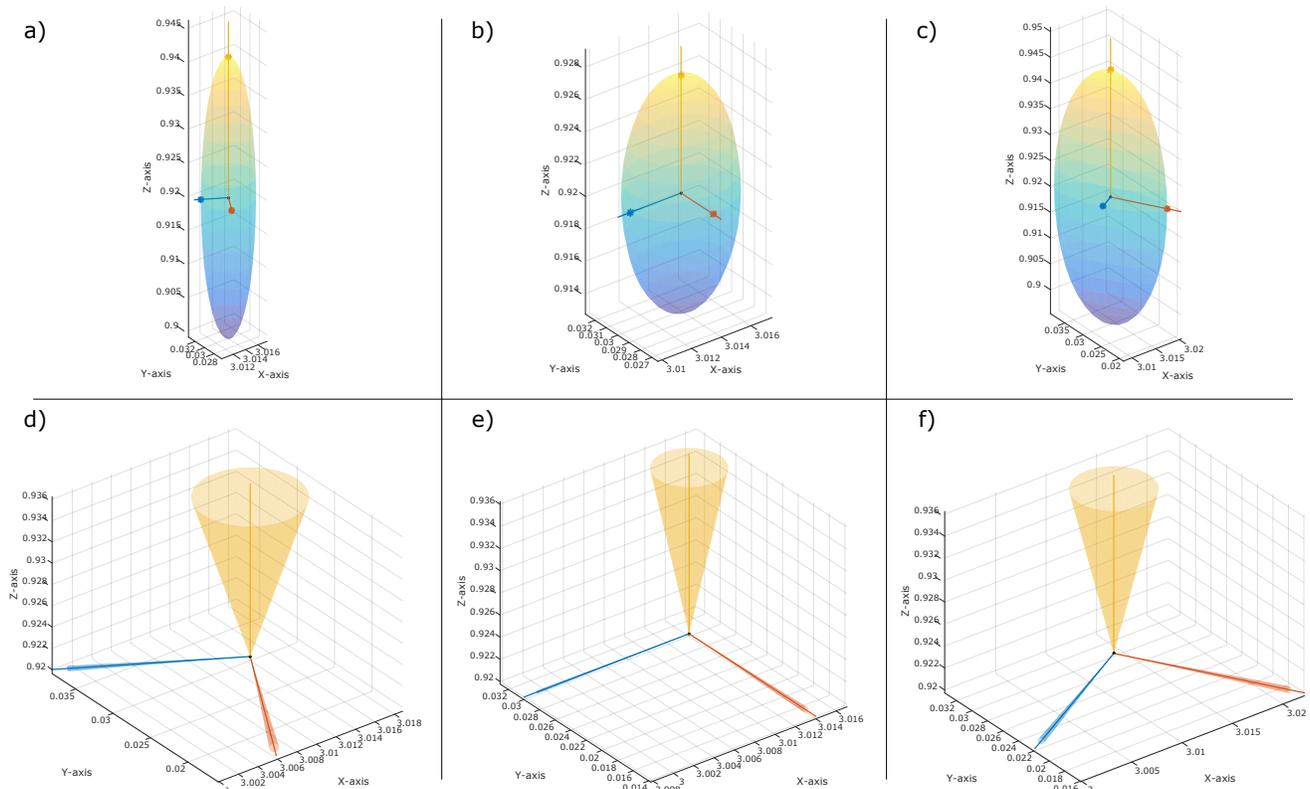


Fig. 10. All of the figures above show the true pose of the mobile receiver node with its position as the center of the axes and its rotation represented by the direction of the axes. a), b) & c) The ellipse that is drawn over these axes represents standard deviation on the position estimate for the given pose. d), e) & f) The cones drawn over each axis represent the standard deviation on the rotation estimate for the given pose with the opening angle of the cone equal to the standard deviation.

the entire mobile receiver node. This optimization in the configuration of the microphone sensor array and the transducer array will result in more stable and less biased pose estimates. Besides the proposed improvements we will use the data of the IMU sensor that has been incorporated into the design, in combination with the ultrasonic measurements using a sensor fusion algorithm to further enhance the accuracy, precision and update rate of the overall pose estimation system.

#### ACKNOWLEDGMENTS

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

#### REFERENCES

- [1] 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>
- [2] 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/>
- [3] L. Xia, C.-C. Chen, and J. K. Aggarwal, "Human detection using depth information by Kinect," in *CVPR 2011 WORKSHOPS*. IEEE, jun 2011, pp. 15–22. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5981811>
- [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/oikonom> <http://www.ics.forth.gr/kyriazis> <http://www.ics.forth.gr/argyros>
- [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://dl.acm.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.
- [9] D. Vlastic, 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>
- [10] 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>
- [11] 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)
- [12] 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>
- [13] D. Laurijssen, S. Truijten, 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.
- [14] J. Steckel, “Sonar System Combining an Emitter Array With a Sparse Receiver Array for Air-Coupled Applications,” *IEEE Sensors Journal*, vol. 15, no. 6, pp. 3446–3452, jun 2015. [Online]. Available: <http://ieeexplore.ieee.org/ielx7/7361/7086408/07008433.pdf>
- [15] Senscomp, “Series 7000 Ultrasonic Sensor Features :,” pp. 1–3, 2014.
- [16] STMicroelectronics, “STM32F407xx Datasheet,” p. 201, 2015.
- [17] Future Technology Devices International Ltd, “FT232H Single Channel Hi-Speed USB to Multipurpose UART/FIFO IC.”
- [18] ST Microelectronics, “LSM9DS0 iNEMO inertial module: 3D accelerometer, 3D gyroscope, 3D magnetometer,” 2013.
- [19] Texas Instruments, “TPS6300x High-Efficient Single Inductor Buck-Boost Converter,” 2015.
- [20] Knowles Electronics, “SPU0410HR5H-PB,” p. 11, 2013.
- [21] Analog Devices, “Precision, 20 MHz, CMOS, Rail-to-Rail Input/Output Operational Amplifiers.”
- [22] Silicon Labs, “Si4464/63/61/60 - High-Performance, Low-Current Transceiver,” 2012.
- [23] DecaWave, “ScenSor DWM1000 Module.”
- [24] Microchip, “MCP73833 - Stand-Alone Linear Li-Ion / Li-Polymer Charge Management Controller.”
- [25] Multicomp, “Lithium-ion Battery.”
- [26] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, ser. Detection, Estimation, and Modulation Theory. Wiley, 2004, no. pt. 1. [Online]. Available: <https://books.google.be/books?id=Xzp7VkuFqXYC>
- [27] B. Fontaine and H. Peremans, “Determining biosonar images using sparse representations,” *The Journal of the Acoustical Society of America*, vol. 125, no. 5, p. 3052, 2009. [Online]. Available: <http://scitation.aip.org/content/asa/journal/jasa/125/5/10.1121/1.3101485>
- [28] A. Jimenez Martin, A. Hernandez Alonso, D. Ruiz, I. Gude, C. De Marziani, M. C. Perez, F. J. Alvarez, C. Gutierrez, and J. Urena, “EMFi-Based Ultrasonic Sensory Array for 3D Localization of Reflectors Using Positioning Algorithms,” *IEEE Sensors Journal*, vol. 15, no. 5, pp. 2951–2962, may 2015. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6994821>
- [29] J. Steckel, A. Boen, D. Vanderelst, and H. Peremans, “EMFit based Ultrasonic Phased Arrays with evolved Weights for Biomimetic Target Localization,” in *ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, Bruges, 2012.