An Ultrasonic Six Degrees-of-Freedom Pose Estimation Sensor — **Source link**

Dennis Laurijssen, Steven Truijen, Wim Saeys, Walter Daems ...+1 more authors

Institutions: University of Antwerp

Topics: Microphone array, Pose, Microphone, Video game and Node (networking)

Related papers:

- Highly Accurate Time-of-Flight Measurement Technique Based on Phase-Correlation for Ultrasonic Ranging

- Accuracy of distance measurements using signal tracking of spread-spectrum ultrasonic waves with CDMA

- Wideband Acoustic Positioning With Precision Calibration and Joint Parameter Estimation

- Indoor Positioning Technologies

- Opto-acoustic distance measurement using spread spectrum techniques and carrier phase measurements

Share this paper:

View more about this paper here: https://typeset.io/papers/an-ultrasonic-six-degrees-of-freedom-pose-estimation-sensor-3f3wyrnn6p

An ultrasonic six degrees-of-freedom pose estimation sensor

An Ultrasonic Six Degrees-of-Freedom Pose Estimation Sensor

Dennis Laurijssen*‡, Steven Truijen†‡, Wim Saeys†, Walter Daems*‡, Jan Steckel*‡

*FTI CoSys Lab, University of Antwerp, Belgium

†FMH REVAKI, University of Antwerp, Belgium

‡Centre for Care Technology, University of Antwerp, Belgium

Abstract—Motion capture and human body pose estimation systems have become a more common appliance nowadays because of the movie and video game industry. These measurement systems have been proven to be useful for other applications besides entertainment. One of these applications is motion analysis, which can be used for improving the form of athletes or for providing an objective validation tool for rehabilitation treatments. These analyses are done using high-accuracy measurement systems which result in high costs. Although there are some consumer products (e.g. the Microsoft Kinect) that offer movement tracking at a low cost, the accuracy does not suffice for clinical movement analysis applications. Our research therefore focuses on reducing the cost of a human body pose estimation system while retaining the required accuracy. The proposed solution comprises of an embedded ultrasonic transmitter and receiver subsystem. The receiver subsystem consists of multiple mobile nodes that are equipped with a small microphone array (at least 3 microphones). Each mobile receiver node captures the encoded simultaneously broadcast ultrasonic transmissions from a distributed transmitter array (which consists of at least 3 elements). Using signal processing, a distance can be calculated between each transmitter and microphone resulting in at least 9 distances for each mobile node. Using these distances in combination with the position of the transmitters and the microphone array configuration, the XYZ-position of the mobile node and its rotation about these axes (6 Degrees-of-Freedom) can be estimated. The combination of low-cost embedded and ultrasonics hardware that forms the transducer and receiver subsystem (consisting of multiple mobile receiver nodes) together with powerful signal processing techniques yields a high-accuracy pose estimation system, which can be used as an affordable tool in various fields and applications (e.g. gait analysis for rehabilitation purposes).

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

I. INTRODUCTION

Articulated body pose estimation or in a more popular term Motion Capture (Mo-Cap) has become an increasingly popular technique that is used to digitally capture the position and rotation of an object or person as a function of time. It can be found in various applications nowadays and its origins can be traced back to rotoscoping [1] which is an animation technique where live-action footage is traced frame per frame by animation artists in order to create an animated film. This technique has been long surpassed by newer technologies and innovations which are now commonly used in the entertainment sector (cinema, television and video games). An application that currently makes use of high-accuracy body

pose estimation systems is gait analysis, which focuses on measuring and analyzing the movements and reciprocal angles between the subject's body parts. The analysis can be used to enhance the technique of athletes in order to increase their performance or to assist physiotherapists objectively validate the rehabilitation therapy of their patients [2], [3]. These high-accuracy systems also come with an overall high cost which makes the technology unviable for most physiotherapists, rehabilitation centers or even hospitals. We wish to reduce the cost for Mo-Cap systems while retaining the desired measurement accuracy in order to increase the availability and viability of high-accuracy body pose estimation applications.

When looking at the different motion capture systems we can distinguish several system topologies. The most known and popular systems rely on optical sensing where multiple cameras are used in combination with active or passive markers (e.g. Vicon [4]–[6]). These markers are fixed at known positions on the object or person and are easily distinguishable in recordings which enables the system to calculate the position of each marker and therefore the entire pose of the person or object. These systems exhibit a high-accuracy but require a great deal of calibration, a fixed setup of multiple cameras in an environment and are inherently costly.

Another popular method for performing motion capture is the use of Inertial Measurement Units (IMUs) such as an accelerometer, gyroscope and magnetometer (digital compass) [7]. One of the leading commercial systems of this type of Mo-Cap systems is MVN Biomech from XSENS [8], [9], which integrates the inertial sensors into nodes that can be inserted into a housing that can be strapped onto the various places on the person or onto an object. Full body suits also exist with the sensor nodes already in place. During the motion tracking, the nodes send the inertial measurement data wirelessly to a gateway device that is connected to the computer that can register and record this information. This technology makes it possible to use motion capture in various environments with high accuracy and without suffering from occlusion. These inertial systems however do suffer from drift [10] that requires recalibration. In order to perform a full body measurement a high number of measure points and therefore sensor nodes is required which in turn makes this an expensive system.

Less frequently used techniques for performing pose estimation are image-based systems that rely on computer vision techniques to analyze video footage [11], magnetic systems

which measure changes in the magnetic fields [12] (e.g. the MotionStar system makes use of Fluxgate magnetometers), mechanical systems (e.g. Meta Motion Gypsy) that use an exo-skeleton to measure the joint angles between the limbs and the various body parts and acoustics-based systems [13]. We propose using an acoustics-based body pose estimation system that uses sound waves in the ultrasonic frequency spectrum to measure the distance between fixed transmitters and mobile receiver nodes based on the Time-of-Flight (ToF) of the emitted signals. Since the required hardware can be composed of low-cost components, a relatively cheap measurement system for pose estimation can be established. When using multiple mobile receiver nodes placed on various body parts this system could be used for human body pose estimation.

In the remainder of this paper the topology of the acoustic pose estimation system will be detailed after which the experimental hardware prototypes of the embedded ultrasonic measurement system are presented. Compared to previous work [14] these prototypes enable us to simultaneously use multiple ultrasonic transducers which are used for preliminary validation of the fundamental ultrasonic ToF-measurements. The addition of more transducers for emitting the ultrasonic sound waves allows us to estimate the full pose of our mobile receiver node prototype with real-life experiments. The results of these pose estimation experiments using the prototypes of the transducer and mobile receiver subsystem are described in the section IV. In the final section, the conclusions of our research will be discussed.

II. System Topology

The embedded ultrasonic measurement system consists of mobile sensor nodes containing small Micro-ElectroMechanical Systems (MEMS) microphones that form a microphone array (at least 3 microphones), which simultaneously receives encoded ultrasonic sound waves emitted from fixed transmitters (at least 3 transmitters) [14], [15]. The complete pose of the mobile sensor (6 Degrees-of-Freedom) is estimated using the ultrasonic range estimates and maximum likelihood estimations. The pose of the sensor consists of the X, Y and Z-position of the sensor in a Cartesian coordinate system and the rotation about the three principle axes. By attaching these mobile sensor nodes to various parts of the human body, a measurement system for human body pose estimation can be established.

III. Experimental Prototype

In order to validate the proposed system topology, an experimental prototype measurement system has been built using both custom designed circuit boards and off-the-shelve ultrasonic components, development kits and modules. The experimental prototype measurement system can be divided into two subsystems, the transducer subsystem, which transmits the ultrasonic signals, and the (mobile) receiver subsystem, which receives these broadcast ultrasonic signals. In order to synchronize and link the subsystems to each other a physical connection has been made but this can be easily replaced by a wireless connection (e.g. a sub-1GHz wireless link established using a CC1101 or CC430 IC).

A. Hardware Setup

Both the transducer and the mobile receiver subsystems have an ARM Cortex M4 microcontroller from STMicroelectronics (STM32F4) as their core. These microcontrollers can be clocked to speeds up to 180 MHz and provide a high number of on-board peripherals, which make these cores very suitable for high-performance embedded applications.

The transducer subsystem circuit board, which is schematically represented in Figure 1a), has been developed with the aforementioned ARM Cortex M4 microcontroller in combination with eight Digital-to-Analog Converters (DACs), more specifically the DAC121S101 from Texas Instruments. These DAC ICs share the same chip select and clock lines, which allows the synchronization of the DACs to the sample level. To provide digital data to the DACs, the Input/Output (I/O) microcontroller lines which connect to the input of each DAC have been specifically chosen to share the same I/O port. By writing a 16-bit value to this port, all of these I/O lines can be simultaneously controlled which in turn enables the system to simultaneously generate up to eight different analog signals. These are fed into an array of eight linear amplifiers which each drive a Senscomp 7000 ultrasonic transducer that is also formally known as the Polaroid ultrasonic transducer. [16].

In order to interface the transducer subsystem circuit board with a computer, it has been fitted with an FT232H, a Hi-Speed Single Channel USB-UART chip. This computer interface is used for initiating transmissions by sending Universal Synchronous Asynchronous Receiver Transmitter (USART) commands to the microcontroller ranging from a simple terminal application to more advanced applications like Matlab. For expansion purposes a number of GPIO headers, which feature communication peripherals, have been accommodated as well as a micro SD card connector for data logging.

Currently the main use of these GPIO pins is linking the transducer board to the receiver subsystem with either a physical or a wireless link. The link is used to notify the receiver that the Polaroid transducers will broadcast their ultrasonic sound waves, which synchronizes the start of measuring the incoming data. This link is therefore crucial for performing the Time-of-Flight measurements and the distance estimation between the transducers and the receiver board. The assembled transducer circuit board is shown in Figure 1b).

The (mobile) receiver subsystem for this experimental hardware setup consists of an STM32F4 Discovery board which incorporates an STM32F407 microcontroller, an on-board programmer, a 3-axes accelerometer, an audio DAC, some LEDs, pushbuttons and a USB On-The-Go (OTG) interface. In order to tune this development board to the requirements of the receiver subsystem, a microphone sensor array board has been used which features one Operational Amplifier (Op-Amp) per microphone channel. The sensor array, which has been developed by Peremans, Steckel, et al. for bio-mimetic and broadband 3D sonar research projects [17], [18], comprises of 32 Knowles Electronics SPU0410HR5H microphones. Using the microphone array board allows for choosing various combinations of three microphones during the experimental validation of the human body pose estimation system. The
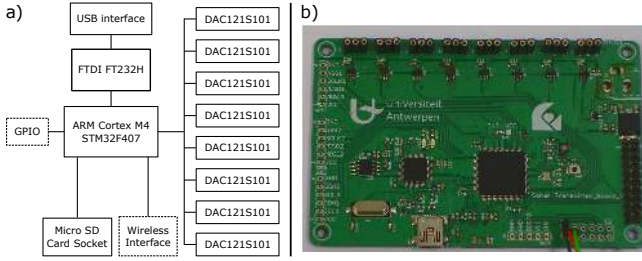
Fig. 1. a) schematic overview of the transmitter circuit board which incorporates an ARM Cortex M4 (STM32F407) microcontroller together with an FTDI FT232 Hi-Speed USB-UART IC, a micro SD card slot and eight Digital-to-Analog (DAC121S101) converters which synchronously generate analog signals which connect to the external linear amplifiers. The transducer board is fitted with GPIO header pins for additional peripherals such as a wireless link module. b) Picture of the developed printed circuit board for the transducer subsystem (10 cm by 6.6 cm).

combination of microphones used in these experiments (indicated in Figure 2b) have been chosen to fit within a rectangle with a given size (5cm by 5cm) and form a non-equilateral triangle. The microphone outputs are each connected to a separate Analog-to-Digital Converter (ADC) channel. This allows the receiver to measure the incoming ultrasonic sound waves simultaneously (in parallel) ensuring that the measured data from the three microphones is in-sync with each other.

Like the transducer subsystem, the receiver features a high number of GPIO headers which are used to interface the microcontroller to a number of peripherals, shown in Figure 2a). Besides connecting the Discovery board to the microphone board, these input and output pins are used for connecting the transducer and receiver board to each other for triggering purposes and providing the subsystem with a communication channel for transferring data to a computer using an FTDI FT232H Hi-Speed Single Channel USB-UART cable. In this case, the fast USB interface is mainly used for transferring the digital microphone samples to a computer in a timely manner. Once these values are transferred, they can be processed in the Matlab environment.
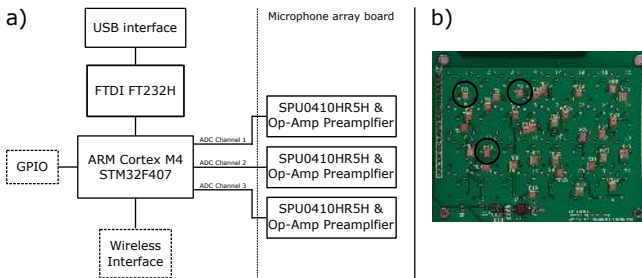

Fig. 2. a) schematic overview of the receiver circuit board which incorporates an ARM Cortex M4 (STM32F407) microcontroller together with an FTDI FT232H Hi-Speed USB-UART IC. An STM32F4 Discovery board has been used for the receiver system which is fitted with all GPIO pins for additional peripherals which can link to a wireless link module. b) A close up picture of the microphone sensor array board which points out which microphones have been used during the experimental hardware setup.

B. Transducer Signaling

Since the goal is to estimate the pose of the mobile receiver nodes during movement, sequential ultrasonic transmissions

for every transducer would yield incoherent distance measurements, resulting in inaccurate estimates. Therefore, all the transducers need to broadcast their ultrasonic pulses at the same time which means that the transmission channel, in this case the air between the transducers and the receivers, has to be shared. In order to combine the three different transmissions in one single transmission channel a spread-spectrum technique is used in combination with Frequency Multiplexing.

The used spread-spectrum technique is implemented by generating three sequences of independent identically distributed random variables also known as white noise. These sequences are individually stored and used as the reference signals for each transducer. Due to the features of white noise it is very suitable to be used for matched filtering [15] which is used for the signal processing at the receiver side. In order to ensure the separation of the transmitter signals and thereby remaining within the transducer's frequency limits, each random variable sequence is filtered using a bandpass filter for each transducer. The frequency ranges between the cutoff frequencies of the bandpass filters have been chosen such that the relative bandwidth $B_r = f_H/f_L$ will be the same. The cutoff frequencies in our experiments are $f_{L_1} = 22\text{kHz}$, $f_{H_1} = 33\text{kHz}$, $f_{L_2} = 33\text{kHz}$, $f_{H_2} = 50\text{kHz}$, $f_{L_3} = 50\text{kHz}$, $f_{H_3} = 75\text{kHz}$, thus ensuring a relative bandwidth of about 1.5 and remaining within the transducer's frequency ranges. By using a fixed frequency ratio, the level of the side-lobes is fairly uniform when using cross-correlation techniques [19]. Figure 3 shows the cross-correlation of the generated transducer base signals. The base signal for a transducer in the figure corresponds to the row number and is cross-correlated with a base signal which corresponds to the column number, e.g. the sub-plot in the lower right corner of the image represents the auto-correlation of the base signal for the third transducer where the left sub-plot on the second row represents the cross-correlation of the base signal for the second transducer with the base signal of the first transducer. These plots attest to the efficacy of the proposed method.

Due to the relatively narrow peaks of the auto-correlation function of the transmitted signals, which can be seen in Figure 3, the system will exhibit a high resistance to reflections of transmitted signals. For our system setup the path of a reflected signal needs to be at least 5.8cm longer than its direct path. This has been demonstrated by Steckel, Boen and Peremans [20] by applying the Rayleigh criterion.

C. Receiver Signal Processing

At the receiver sub-system the incoming ultrasonic sound waves are recorded using the three selected microphones on the microphone array sensor board. These microphones are connected to three individual Analog-to-Digital Converters (ADCs), which are configured to work simultaneously (in parallel) ensuring that the recorded signals are in sync with each other. Once the transducer sub-system broadcasts its transmission, the receiver is triggered using a wired connection to start the ADCs with a sampling frequency of 360kHz (which is well beyond the Nyquist rate) for a predefined number of samples. The current number of samples allows
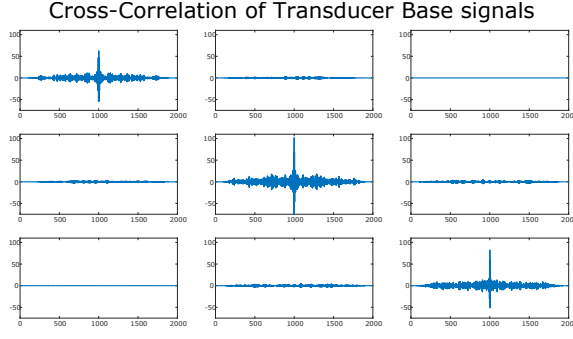
Fig. 3. Cross-correlations between the three broadcast signals. Each row corresponds to a transducer base signal which is cross-correlated to a base signal which corresponds to the column number. E.g. the sub-plot in the center of the first row is cross-correlation between the base signal of the first transducer with the base signal of the second transducer. The main diagonal sub-plots represent the cross-correlation function of the transducer base signal with itself i.e. the auto-correlation function.

capturing an entire ultrasonic emission from a transducer at a distance of 5.9m away from the microphones. Once all of the samples have been recorded these values are transferred to a computer using the serial interface.

Each of these three microphone signals $s_{m_k}(t)$ consists of a combination of three emitted sound waves. Using cross-correlation techniques the time of arrival of each individual recorded signal can be determined. For this purpose, a frequency-weighted generalized cross-correlation (GCC) function [21], [22] has been used with the pre-recorded transducer base signals $s_{b_i}(t)$ and the three microphone signals $s_{m_k}(t)$ as its inputs:

$$s_{r_{k,i}} = \mathcal{F}^{-1} \left\{ \frac{S_{m_k}(j\omega)S_{b_i}(j\omega)}{|S_{m_k}(j\omega)||S_{b_i}(j\omega)|} . \mathcal{W}_i(j\omega) \right\} \quad (1)$$

with $s_{r_{k,i}}$ the resulting signal of the frequency-weighted generalized cross-correlation function with indexes k, i as the used microphone and base signal, $S_{m_k}(j\omega)$ the Fourier transformation of a microphone signal $s_{m_k}(t)$ with index k, $S_{b_i}(j\omega)$ as the Fourier transformation of the pre-recorded transducer base signal $s_{b_i}(t)$ with index i and $\mathcal{W}$ as the weight function with:

$$\mathcal{W}_i(j\omega) = \begin{cases} 1 & \forall \quad f_{L_i} < \omega < f_{H_i} \\ 0 & \forall \quad \omega < f_{L_i} \\ 0 & \forall \quad \omega > f_{H_i} \end{cases} \quad (2)$$

Figure 4a) shows a simulated microphone input signal $s_m(t)$ in the time domain which comprises of a sequential combination of three different band-limited white noise sequences. Plots b), c) and d) of Figure 4 show the output of using the frequency-weighted generalized cross-correlation function as described above with the simulated microphone signal as the input with respectively the transducer base signals $s_{b_1}(t), s_{b_2}(t), s_{b_3}(t)$. The peaks of the output signals display the start of each signal in the microphone input signal $s_m(t)$, thus representing the Time of Arrival. In the aforementioned simulation the transducers broadcast their signals sequentially, Plot 5a) represents a simulated microphone input signal $s_m(t)$ that correspond to a more realistic input signal where the transducers would have broadcast simultaneously. Figure 5b)
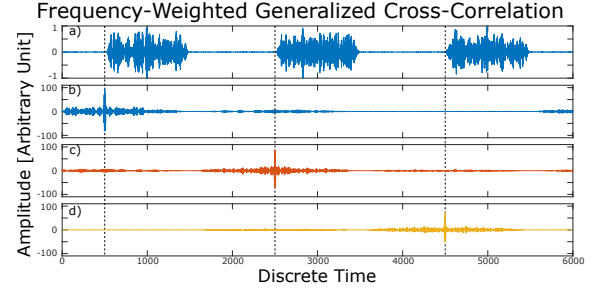

Fig. 4. a) Simulated microphone signal where the transducers emit sequentially with noise added. b, c and d) The result of applying a Frequency-Weighted Generalized Cross-Correlation function on the simulated microphone signal with three pre-recorded transducer base signals. Each peak of a plotted cross-correlation result represents the Time of Arrival of the emitted signal by a transducer.

shows the three outputs of the frequency-weighted generalized cross-correlation function where the peaks indicate the Time of Arrival of each signal. When applying this technique to
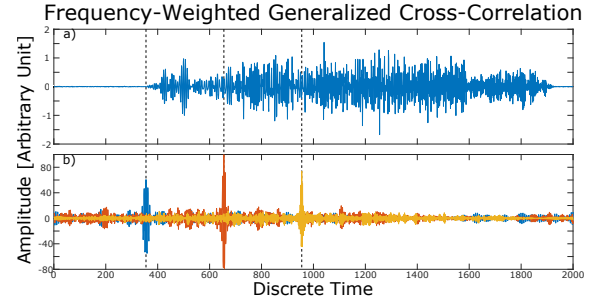

Fig. 5. a) Shows a simulated microphone signal where the transducers emit simultaneously with added noise. b) Demonstrates the results of applying a Frequency-Weighted Generalized Cross-Correlation function on the simulated microphone signal with three pre-recorded transducer base signals. Each peak (blue, red and yellow) of the plotted cross-correlation result represents the Time of Arrival of the emitted signal by a transducer.

the three individual microphone signals $s_{m_k}(t)$ with the three pre-recorded transducer base signals $s_{b_i}(t)$ it is possible to calculate the distances between each transducer and each microphone using the speed of sound in air thus resulting in the measurement vector $\overrightarrow{M}$ with $d_{k,i}$ the distances for each microphone-transducer pair:

$$\overrightarrow{M} = [d_{1,1}, d_{1,2}, ..., d_{k,i}]^T \quad (3)$$

In order to calculate the measurement vector $\overrightarrow{M}$ precisely the speed of sound in air needs to be approximated. Since the speed of sound in air is affected by the altitude and the air's temperature and humidity, these factors need to be taken into account. Currently only the room temperature has been taken into account using the equation for the speed of sound in dry air (Equation 4) for calculating $\overrightarrow{M}$, with T_C the temperature in degrees Celsius.

$$c_{air} = (331.3 + 0.606.T_C) \text{ m/s} \quad (4)$$

D. Pose Estimation

Using the measurement vector $\overrightarrow{M}$ which comprises the nine distances between each microphone - transducer pair,

the transducer coordinates (X, Y, Z) and the microphone array configuration the full pose $\overrightarrow{P_{Est}}$ $(X, Y, Z$ and $\alpha, \beta, \gamma)$ for a mobile node can be estimated with a maximum likelihood estimator in combination with a multidimensional unconstrained nonlinear minimization function. The nonlinear minimization function will generate a full pose $\overrightarrow{P_i}$ for a mobile node with every iteration i which is converted into a distance vector $\overrightarrow{D}$ which is calculated using the given transducer coordinates and the microphone array configuration. Using the measurement vector $\overrightarrow{M}$ and distance vector $\overrightarrow{D}$, the likelihood of the generated mobile node pose can be calculated using:

$$\mathcal{L}(\overrightarrow{M}|\overrightarrow{P_i}) = \frac{1}{(2\pi)^{\frac{1}{2}}|\Sigma|^{\frac{1}{2}}} \cdot \exp\left[-\frac{1}{2}(\overrightarrow{M} - \overrightarrow{D})^T \Sigma^{-1}(\overrightarrow{M} - \overrightarrow{D})\right] \tag{5}$$

with Σ the covariance matrix of the distance measurements of the microphone - transducer pairs. In order to use the calculated likelihood $\mathcal{L}(\overrightarrow{M}|\overrightarrow{P_i})$ as an evaluation criterion for the nonlinear minimization function, the natural logarithm of the likelihood $\log \mathcal{L}(\overrightarrow{M}|\overrightarrow{P_i})$ is used. The resulting log-likelihood is used to determine whether the generated pose for the mobile node fits the measured distances between the microphone-transducer pairs. The nonlinear minimization function will minimize the log-likelihood with every iteration thus fitting the generated pose $\overrightarrow{P_i}$ to the measurement vector $\overrightarrow{M}$ until a pre-defined stopping criterion is reached.

E. Experimental Setup

Using the hardware prototype of the transducer- and receiver subsystems in combination with the described techniques in the previous sections, an experimental setup, which is shown in Figure 7a), has been created for testing the efficacy of our proposed pose estimation system prototypes. In this experimental setup the prototype transducer subsystem is connected to three Senscomp 7000 ultrasonic transducers which have been placed on known positions in the environment. The placement of these transducers in the environment affects the pose estimation accuracy. When more variation is introduced in the X, Y, Z coordinates of the different transducers there will be less ambiguity in the calculated distances between each transducer - microphone pair which will benefit the accuracy of pose estimations of the mobile nodes. Therefore the transducers shown in Figure 7a) are positioned to form a non-equilateral triangle in a non-horizontal plane, the transducer coordinates in the environment are shown in Table I. The receiver subsystem currently comprises of only one prototype of a mobile receiver node for this setup and has been placed in the environment at a position $(X = 3.0\text{m}, Y = -0.018\text{m}, Z = 0.84\text{m})$ and orientation $(\alpha = 0.0°, \beta = 0.0°, \gamma = 180.0°)$ which has to be determined using our proposed hardware and signal processing schemes.

TABLE I
TRANSDUCER COORDINATES IN THE ENVIRONMENT

	X	Y	Z
Transducer 1	1.0m	1.5m	1.0m
Transducer 2	0.0m	0.0m	0.82m
Transducer 3	0.5m	-1.0m	1.70m

TABLE II
REFERENCE DISTANCES BETWEEN EACH TRANSDUCER - MICROPHONE PAIR

	Transducer 1	Transducer 2	Transducer 3
Microphone 1	2.555m	3.003m	2.775m
Microphone 2	2.564m	3.004m	2.762m
Microphone 3	2.548m	3.003m	2.771m

IV. RESULTS

Using the experimental system setup the aforementioned techniques for estimating the pose of the mobile receiver node have been validated. In order to perform the actual pose estimation, the operation and accuracy of the proposed frequency-weighted generalized cross-correlation technique needs to be evaluated in order to accurately determine the Time-of-Flight for every transducer - microphone combination, thus creating the distance vector $\overrightarrow{D}$.

A. ToF Measurements

Using the experimental setup, the three different band-limited pseudo random sequences for each Senscomp transducer were first individually recorded in order to be used as the reference signal for the corresponding transducer. Once every transducer reference signal was recorded, the transducers were simultaneously actuated by the transducer subsystem which in turn triggered the mobile receiver subsystem to start recording the incoming ultrasonic sound waves by enabling the three synchronous ADC channels which are connected to the microphone array board. Once all the samples are collected they are transmitted to a computer using the high-speed UART-to-USB connection. Once received on the computer side, these samples are used to reconstruct the microphone signals which are fed into the proposed frequency-weighted generalized cross-correlation filter together with the transducer reference signals. The outcome of this filter is shown in Figure 8 in which each subplot represents a microphone signal which is used as filter input in combination with one of the three reference signals. Every subplot therefore includes three filter outputs where the peak of the output signal represents the arrival time for the emitted transducer sound waves. Using the arrival time in combination with the speed of sound in air, the distance vector $\overrightarrow{D}$ is calculated. For verifying these calculated distances, the distances between each transducer - microphone pair have been measured using a Bosch GLM 150 laser distance estimator to serve as the reference, which are shown in Table II, after which the experimental hardware is used to measure the distances 100 times. For every measurement, the difference between the reference and the measurement iteration is calculated and stored as the error on the distance measurement. After 100 measurement iterations the mean of these errors is calculated, these results can be found as the Mean Absolute Error (MAE) and as the mean error percentage in Table III. The standard deviation (σ) on the distance measurement error is expressed in mm and percentage, which can be found in Table IV.
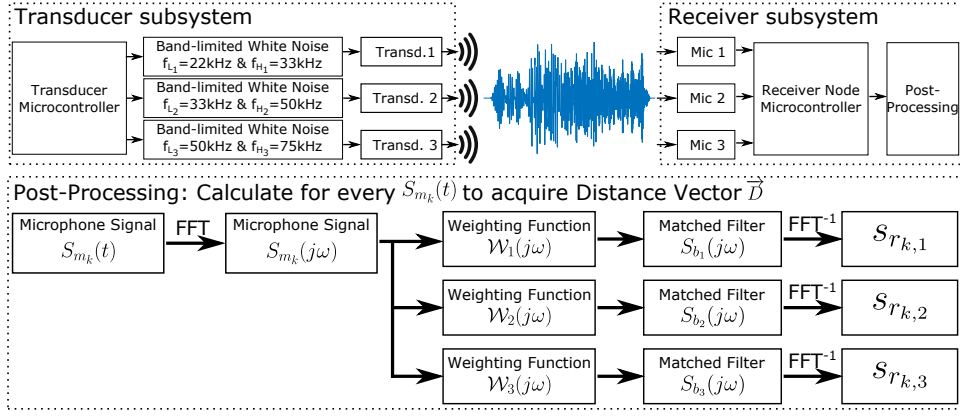
Fig. 6. Flowchart which demonstrates how the pose estimation system calculates the distances (distance vector $\vec{D}$) between the transducers and the microphones of the mobile receiver node.
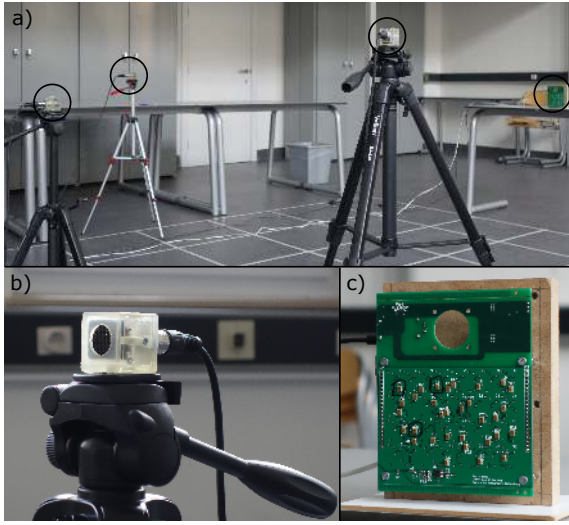

Fig. 7. a) Shows the hardware setup which has been used for conducting the experiments. Three transducers, enclosed in a 3D-printed housing which are be mounted on a standard camera tripod, have been set up in the environment on known positions. In order to have more *unique* distance measurements between every microphone - transducer pair, the transducers have been placed at different X, Y, Z coordinates in the environment. The circled transducer and microphone array board are respectively shown in close-up in b) and c).
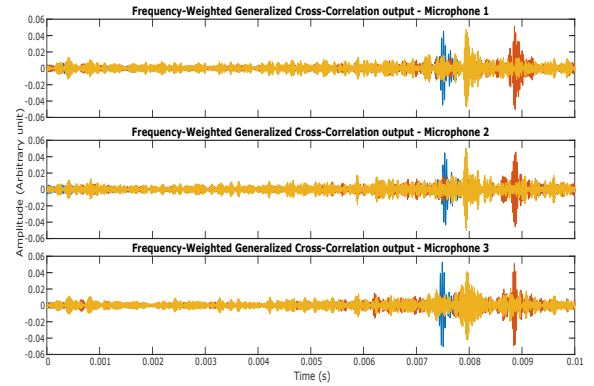

Fig. 8. The output of the proposed frequency-weighted generalized cross-correlation filter is plotted in this figure. Every subplot represents one of the three microphone signals which is used in combination with the three reference signals for every transducer. The output of the frequency-weighted generalized cross-correlation filter with every microphone - transducer combination is therefore shown in this figure. Each peak in the output signals represent the arrival time of the emitted signal originating from a transducer to reach the given microphone. Using these arrival times in combination with the sound of speed in air, the distance vector $\vec{D}$ is calculated.

TABLE III
MEAN ERROR ON DISTANCE MEASUREMENTS

	Transducer 1		Transducer 2		Transducer 3	
	MAE	%	MAE	%	MAE	%
Microphone 1	1.3mm	0.005%	18.8mm	0.0063%	23.9mm	0.0086%
Microphone 2	8.7mm	0.0034%	12.8mm	0.0043%	25.6mm	0.0093%
Microphone 3	16.5mm	0.0065%	19.5mm	0.0065%	30.7mm	0.0111%

TABLE IV
STANDARD DEVIATION ON DISTANCE MEASUREMENTS

	Transducer 1		Transducer 2		Transducer 3	
	σ	%	σ	%	σ	%
Microphone 1	1.7mm	0.0007%	1.9mm	0.0006%	2.2mm	0.0008
Microphone 2	1.7mm	0.0007%	2.9mm	0.0010%	2.9mm	0.0011%
Microphone 3	1.7mm	0.0007%	1.9mm	0.0006%	1.7mm	0.0006%

B. Static Pose Estimation

Although the experimental system currently does not meet the required accuracy and precision of the most demanding motion capture purposes, the measurement iterations that were collected during the ToF-measurements were used for estimating the full pose $\overrightarrow{P_{Est}}$ of the mobile receiver node. Using the proposed maximum likelihood estimator in combination with a multidimensional unconstrained nonlinear minimization function, with the collected data and system configuration (transducer coordinates and the microphone array configuration) as the input variables, the full pose $\overrightarrow{P_{Est}}$ is estimated over 100 measurement iterations. Table V features the results of these pose estimations, both the mean absolute error and the standard deviation are presented. Figure 9 demonstrates the results of these pose estimations as two three-dimensional plots in which a) shows the reference X, Y, Z-coordinates of the mobile receiver node as the center of the axes. The ellipsoid that is drawn over these axes represents the standard deviation over all the pose estimation iterations with the size of the semi-principal axes equal to the standard deviation for the corresponding axis. Figure 9b) shows the standard deviation on the rotations α, β, γ over all the pose estimations with the

center of the axis plotted at the reference coordinates of the mobile receiver node. The cones that are drawn over these axes have an opening angle that corresponds to the standard deviation over all the pose estimations for the rotation about its axis.

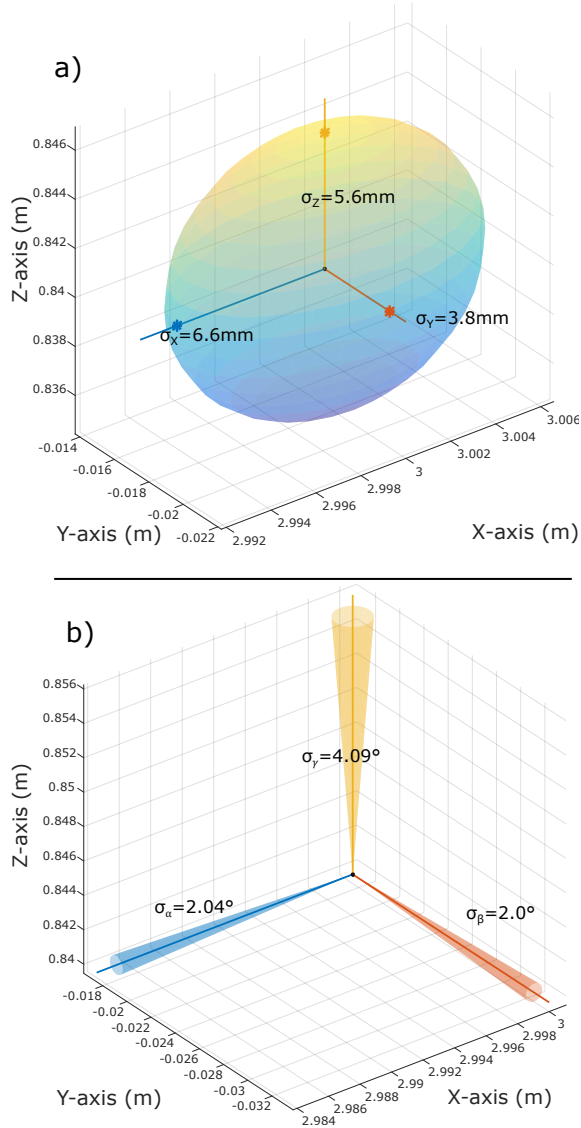|	X	Y	Z	α	β	γ
Mean Error	5.5mm	19.9mm	0.2mm	1.97°	0.23°	3.56°
Standard Deviation	6.6mm	3.8mm	5.6mm	2.04°	1.99°	4.09°

Fig. 9. When estimating the pose of the mobile receiver node 100 times, the estimated pose is fairly stable. The figures both represent the true reference pose of the receiver node drawn in a Cartesian coordinate system with respect to its true position and rotation. In a) the ellipsoid which is drawn over the true reference pose represents the standard deviation on the estimated X, Y, Z positions with the size of the semi-principal axes equal to the standard deviation for the corresponding axis. In b) the standard deviation on the estimated α, β, γ rotations are represented by cones along the corresponding axis with the radius equal to the standard deviation of the corresponding rotation.

C. Pose Estimation of Trajectory

Using Matlab a simulation environment has been developed in which the transducer and microphone array configuration of the mobile receiver node can be declared as one of its variables in combination with the other parameters such as the environment size, temperature, etc. The pose of mobile receiver node is declared at various intervals to create a trajectory through the environment for which the system has to estimate the pose of the simulated mobile node. Using the proposed methods and techniques, the simulation environment is able to calculate a realistic distance vector $\overrightarrow{D}$. In order to produce the full pose of the mobile receiver node, the distance vector $\overrightarrow{D}$ is used as the input for maximum likelihood estimator in combination with the multidimensional unconstrained nonlinear minimization function. Figure 10 shows the simulated environment that corresponds to our system setup in which a) plots the true poses of the simulated mobile node where b) plots the estimated poses of the simulated mobile node.
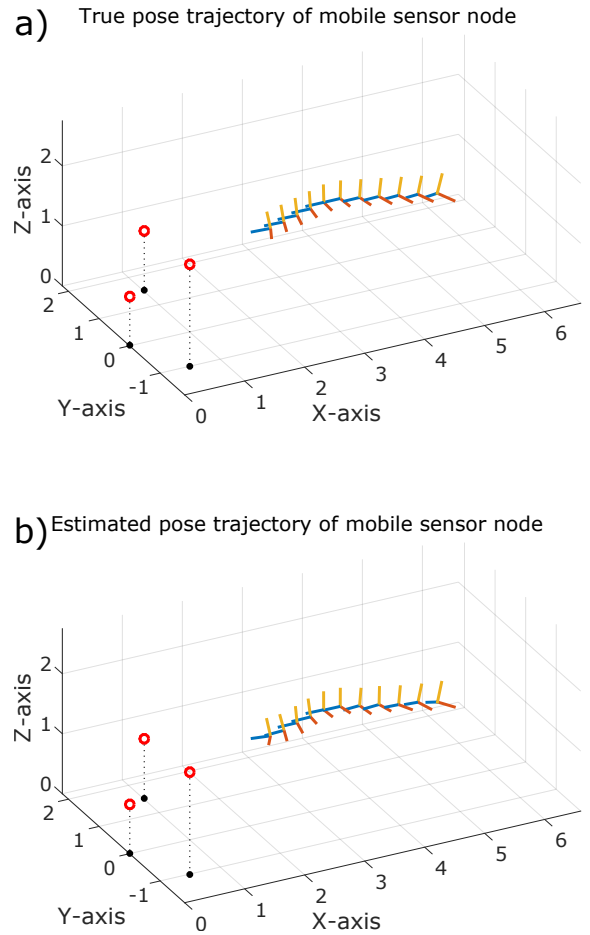

Fig. 10. a) The true trajectory that consist of the true poses of the mobile receiver node generated in the simulation environment. b) Estimation of a trajectory of poses in a simulated environment. Sensor data was generated using a custom-built acoustic simulator. In both a) and b) the red circles represent the transducers which are fixed in the simulation environment.

V. Discussion, Conclusions and Future Work

In this paper, an ultrasonic six Degrees-of-Freedom pose estimation system has been presented which uses three fixed transmitters placed at known positions in the environment and a three-element microphone array on the receiver. Using these peripherals in combination with our embedded hardware prototypes an experimental setup was created which is capable of estimating the X, Y, Z-coordinates and the α, β and γ rotations of a single mobile receiver node. When comparing these results to other pose estimation systems [23], [24], we can state that the results of our proposed pose estimation system features an acceptable accuracy and precision for the X, Y, Z-coordinates, yet lacks the required α, β and γ rotation estimation capabilities for the most demanding motion capture purposes. The current experimental hardware setup will however be improved for our ongoing research into pose estimation systems. One of the main future enhancements of the mobile receiver nodes will be a custom hardware design which integrates an IMU-sensor that would provide the system with a 3-axis gyroscope, accelerometer and magnetometer. These IMU sensor readings can be combined with the ultrasonic measurements within a sensor fusion algorithm that will further enhance the precision and accuracy of the pose estimation capabilities as well as the update rate of the pose estimations. [25] The ultrasonic measurements will in turn be used to limit the effects of the integration drift that is an inherent problem of IMU-sensors [10]. Besides the addition of an IMU sensor, the new hardware design of the mobile receiver node will aim at a small form factor (5cm by 5cm) and integrated wireless connectivity. These features will allow us to create a network of mobile receiver nodes in order to perform human body pose estimation experiments. In addition to an improved mobile receiver node, the Senscomp transducer could be replaced with an Emfit-based [26], [27] transducer for transmitting the ultrasonic band-limited pseudo random sequences. During the conducted experiments the Senscomp transducer suffered from non-linear effects on the output signals and resonance at the 50kHz frequency region. This unwanted effect interferes with the operation of the proposed frequency-weighted generalized cross-correlation technique resulting in less accurate ToF-measurements. Due to the linear properties of an Emfit-based transducer, it is expected to improve and thus enhance the overall efficacy of the pose estimation system. Besides hardware improvements on the proposed pose estimation system, the position configuration of the hardware components in the environment is a crucial component for the efficacy, accuracy and precision. For the transducers this configuration comprises the X, Y, Z-positions in the environment whereas for the receiver it comprises the relative coordinates on the printed circuit board for the microphone array. These configurations can be optimized in order to attain the required accuracy and precision within given constraints (e.g. system cost, size of mobile receiver nodes, size of environment, hardware limitations).

Acknowledgments

References

[1] D. J. Sturman, "A brief history of motion capture for computer character animation," *SIGGRAPH 94, Character Motion Systems, Course notes*, vol. 1, 1994.

[2] W. Saeys, *Postural Control in People with Stroke: Aspects of Evaluation and Treatment : Proefschrift*, ser. Proefschriften UA-GGW. Universiteit Antwerpen, Faculteit Geneeskunde en Gezondheidswetenschappen, 2012. [Online]. Available: https://books.google.be/books?id=_n8UrgEACAAJ

[3] Z. Zhang, Q. Fang, and X. Gu, "Objective Assessment of Upper Limb Mobility for Post-stroke Rehabilitation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 1–1, 2015. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7244204

[4] 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

[5] J. Veneman, E. V. Asseldonk, R. Ekkelenkamp, F. V. D. Helm, H. V. D. Kooij, E. van Asseldonk, F. van der Helm, and H. van der Kooij, "Evaluation of the effect on walking of balance-related degrees of freedom in a robotic gait training device," in *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*, vol. 00, no. c, 2007, pp. 868–875.

[6] B. Birch, E. Haslam, I. Heerah, N. Dechev, and E. J. Park, "Design of a continuous passive and active motion device for hand rehabilitation." *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, vol. 2008, pp. 4306–4309, 2008.

[7] Y. Zhang, K. Chen, J. Yi, T. Liu, and Q. Pan, "Whole-Body Pose Estimation in Human Bicycle Riding Using a Small Set of Wearable Sensors," *IEEE/ASME Transactions on Mechatronics*, pp. 1–1, 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7296666/

[8] M. Cognolato, "Experimental validation of Xsens inertial sensors during clinical and sport motion capture applications," Ph.D. dissertation, 2012. [Online]. Available: http://tesi.cab.unipd.it/40443/1/Tesi_Completa.pdf

[9] J.-T. Zhang, A. C. Novak, B. Brouwer, and Q. Li, "Concurrent validation of Xsens MVN measurement of lower limb joint angular kinematics," *Physiological Measurement*, vol. 34, no. 8, pp. N63–N69, aug 2013. [Online]. Available: http://stacks.iop.org/0967-3334/34/i=8/a=N63?key=crossref.b8cddbd89301a6c8dae28bde2e307266

[10] R. G. J. Damgrave and D. Lutters, "The Drift of the Xsens Moven Motion Capturing Suit during Common Movements in a Working Environment," 2009. [Online]. Available: http://hdl.handle.net/1826/3722

[11] D. Regazzoni, C. Rizzi, C. Comotti, and F. Massa, "Towards Automatic Gait Assessment by Means of RGB-D Mocap," in *Volume 1A: 35th Computers and Information in Engineering Conference*. ASME, aug 2015, p. V01AT02A057.

[12] J. Huang, T. Mori, K. Takashima, S. Hashi, and Y. Kitamura, "IM6D," *ACM Transactions on Graphics*, vol. 34, no. 6, pp. 1–10, oct 2015. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2816795.2818135

[13] 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

[14] 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.

[15] 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

[16] C. Biber, S. Ellin, E. Shenk, and J. Stempeck, "The Polaroid Ultrasonic Ranging System," in *Audio Engineering Society Convention 67*, 1980. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=3680

[17] J. Steckel, A. Boen, and H. Peremans, "Broadband 3-D sonar system using a sparse array for indoor navigation," *IEEE Transactions on Robotics*, vol. 29, pp. 161–171, 2013.

[18] J. Steckel and H. Peremans, "Spatial sampling strategy for a 3D sonar sensor supporting BatSLAM," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, sep 2015, pp. 723–728. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7353452

[19] 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

[20] J. Steckel, A. Boen, and H. Peremans, "Broadband 3-D Sonar System Using a Sparse Array for Indoor Navigation," *IEEE Transactions on Robotics*, vol. 29, no. 1, pp. 161–171, feb 2013.

[21] L. Guan, F. Gu, B. M. Fazenda, A. Ball, Y. Yichun, and T. Pengxiao, "DOA estimation in car for abnormal sound localisation," in *Proceedings of 8th International Conference on Damage Assessment of Structures*, 2009.

[22] F. Grondin and F. Michaud, "Time Difference of Arrival Estimation based on Binary Frequency Mask for Sound Source Localization on Mobile Robots," 2015.

[23] J. Huang, T. Mori, K. Takashima, S. Hashi, and Y. Kitamura, "IM6D: Magnetic Tracking System with 6-DOF Passive Markers for Dexterous 3D Interaction and Motion," *ACM Trans. Graph.*, vol. 34, no. 6, pp. 217:1—-217:10, oct 2015. [Online]. Available: http://doi.acm.org/10.1145/2816795.2818135

[24] H. Tanaka, Y. Sumi, and Y. Matsumoto, "A portable 6-DOF motion tracker using high-accuracy AR markers - First report on the feasibility," in *2015 14th IAPR International Conference on Machine Vision Applications (MVA)*. IEEE, may 2015, pp. 563–566. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7153255

[25] C. He, P. Kazanzides, H. Sen, S. Kim, and Y. Liu, "An Inertial and Optical Sensor Fusion Approach for Six Degree-of-Freedom Pose Estimation," *Sensors*, vol. 15, no. 7, pp. 16 448–16 465, jul 2015. [Online]. Available: http://www.mdpi.com/1424-8220/15/7/16448/

[26] 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

[27] A. Jiménez, Á. Hernández, J. Ureña, M. C. Pérez, F. J. Álvarez, C. De Marziani, J. J. García, and J. M. Villadangos, "EMFi-based ultrasonic transducer for robotics applications," *Sensors and Actuators A: Physical*, vol. 148, no. 1, pp. 342–349, 2008.

Steven Truijen As chairman of the department Rehabilitation Sciences and Physiotherapy (REVAKI), Steven Truijen, PhD, M.Sc., works actively at the development and integration of the department REVAKI in the Faculty of Medicine and Health Sciences at the University of Antwerp. The focus of his research is to add value in health care by multidisciplinary cooperation as a core member of StatUA, the core facility for Statistics of the University of Antwerp, member of the Centre for Care Technology of the University of Antwerp (CZT) and co-promoter of the Multidisciplinary Motor Centre Antwerp (M^{2}OCEAN).

Wim Saeys Dr. Wim Saeys was born in 1982 in Brasschaat and received the M.Sc. degree in Physiotherapy and Rehabilitation Sciences in 2005 at the University College of Antwerp, Belgium. In 2006, he received the M.Sc. degree in neurological rehabilitation at the University of Brussels, Belgium. In the following years, he conducted research on the topic of gait analysis in combination with balance and posture control in people suffering from brain damage. He acquired his PhD in 2012 at the University of Antwerp, Belgium. Currently, he is affiliated with the rehabilitation hospital Revarte and the University of Antwerp.

Walter Daems Walter Daems (S'97-M'02) was born in Wilrijk, Belgium, in 1971. He received the B.Sc. degree in electronics from the Katholieke Industrile Hogeschool Antwerpen, Hoboken, Belgium, in 1994, and the M.Sc. degree in electrical engineering and the Ph.D. degree from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1996 and, 2002, respectively. After a Postdoctoral Fellowship of the Fund for Scientific Research, Flanders, Belgium, he cofounded Kimotion Technologies, Inc. In 2006, he joined the University College Karel de Grote in Hoboken, Belgium, where he was appointed vice dean of academic affairs. In 2013 he was one of the main founders of the faculty of Applied Engineering at the University of Antwerp, Belgium. Since then, he is an associate professor at the faculty of Applied Engineering and chairman of the facultys educational board. He is a member of staff of the Constrained Systems Lab and he is also active in the Centre for Care Technology of the University of Antwerp. His research interests are in digital signal and image processing and the implementation of these systems in embedded technology for application in industry and (health) care.

Dennis Laurijssen Dennis Laurijssen received the M.Sc. degree in electrical engineering in 2012 from the Artesis Hogeschool in Belgium. After getting his M.Sc. degree, he joined the CoSys Lab group in 2012 for a brief period as a researcher. In 2013 he pursued an opportunity at a startup company in San Francisco for a short internship after which he started to work for the Centre for Care Technology at the University of Antwerp until the end of 2015. He is currently working toward a Ph.D. degree within the CoSys Lab group. His research interests include in-air ultrasonic systems, pose estimation systems, wireless communication, embedded (DSP) systems and array signal processing.

Jan Steckel Prof. Dr. Ing Jan Steckel graduated in electronic engineering at the University College Karel de Grote in Hoboken in 2007. In 2012 he obtained his doctoral degree at the University of Antwerp at the Active Perception Lab, with a dissertation titled Array processing for in-air sonar systems drawing inspirations from biology. During this period he developed state-of-the art sonar sensors, both biomimetic and sensor-array based. During his post-doc period he was an active member of the Centre for Care Technology at the University of Antwerp where he was in charge of various healthcare-related projects concerning novel sensor technologies. Furthermore, he pursued industrial exploitation of the patented 3D array sonar sensor which was developed in collaboration during his PhD. In 2015 he became a tenure track professor at the University of Antwerp in the Constrained Systems Lab where he researches sensors, sensor arrays and signal processing algorithms using an embedded, constrained systems approach.