HRTF Measurement by Means of Unsupervised Head Movements With Respect to a Single Fixed Speaker

JONAS REIJNERS, BART PARTOENS, JAN STECKEL, AND HERBERT PEREMANS

1Department of Engineering Management, Universiteit Antwerpen, 2000 Antwerp, Belgium
2Department of Physics, Universiteit Antwerpen, 2020 Antwerpen, Belgium
3FTI CoSys Lab, Universiteit Antwerpen, 2020 Antwerp, Belgium
4Flanders Make Strategic Research Centre, 3920 Lommel, Belgium

Corresponding author: Jonas Reijniers (jonas.reijniers@uantwerpen.be)

This work was supported in part by the Research Foundation Flanders (FWO) under Grant G023619N, and in part by the Agency for Innovation and Entrepreneurship (VLAIO).

ABSTRACT In a standard state-of-the-art measurement the head-related transfer function (HRTF) is obtained in an anechoic room with an elaborate setup involving multiple calibrated loudspeakers. In search for a simplified method that would open up the possibility for an HRTF measurement in a home environment, it has been suggested that this setup could be replaced with one with a single, fixed loudspeaker. In such a setup, the subject samples different directions by moving the head with respect to this loudspeaker, while the head movements are tracked in some way. In this paper, the feasibility of such an approach is studied. To this end, the HRTF is measured in an unmodified (non-anechoic) room by means of a single external speaker and a high resolution head tracking system. The differences between the dynamically obtained HRTF and the standard static HRTF are investigated, and are shown to be mostly due to variable torso reflections.

INDEX TERMS Acoustic measurement, acoustic signal processing, spectral shape, psychoacoustics, acoustical engineering.

I. INTRODUCTION

The HRTF is the acoustic filtering that is performed by the torso, the head and the ears on the incoming sound, before it enters the ear canal. This filtering is different depending on the direction from where the sound originated, and as such, it carries the acoustic cues from which the listener can deduce the location of the sound source [1]. This filtering entails both frequency dependent timing delays (phase), resulting in differences in time of arrival for both ears, and spectral filtering, i.e., to what extent some frequencies are suppressed or amplified.

Since everyone has a different ear, head and torso morphology, everyone has a different HRTF. There has been a lot of research on the individual differences in the HRTF, resulting in numerous HRTF databases that can be accessed online [2]–[4]. The impact of a non-individual HRTFs on three-dimensional (3D) audio perception is often studied by considering a generic HRTF, measured on a dummy head, e.g. KEMAR [5] or Neumann KU100. Generally, it is accepted that an individualized HRTF is superior to a non-individual HRTF, especially regarding sound localization [6], [7], although for some applications a generic HRTF may suffice [4], [8].

Traditionally, an individualized HRTF is measured using specialized and costly infrastructure [2]. Equipped with two in-ear microphones at the entrance of the ear canal, the subject is seated in the middle of an anechoic room. A sound source is then moved on a sphere with the subject’s head in the centre, and for a discrete set of directions a sound is emitted, which is then picked up by the two in-ear microphones. Relating the recorded signals for the sampled directions to the emitted signal allows one to determine the individualized HRTF in great detail. Although new HRTF measurement methods have been developed to simplify and speed up this measurement process [9], such measurements still require costly infrastructure, which makes it impossible to apply them on a large scale.

Individualized HRTFs are generally considered to be of crucial importance for realistic 3D audio reproduction through headphones, and consequently for its breakthrough as an emerging technology. So the question remains: how to acquire individualized HRTFs on a large scale, allowing this...
technology to become accessible for the general public? Over the years, different routes have been taken to individualize the HRTF in a home environment [10], albeit with limited success. The most common alternative strategy is to measure the morphology of torso, head and/or ears and derive an individualized HRTF from these morphological features. This is done either by linking morphological features to a database and exploiting the correlations between morphology and HRTF [11], or by performing numerical simulations of the sound field around a detailed 3D morphology model of the ears, head (and torso) [12].

This detour via morphology is taken to avoid an acoustic measurement, which is deemed impossible or impractical in a home environment. Yet, in recent years it has been suggested that such an acoustic measurement of the individualized HRTF at home could be feasible, using a dynamic approach [13]–[16].

These methods build further on the SNAPSHOT measurement system [17], which was a first attempt to measure an HRTF in a home environment. Here, the subject was seated on a swivel chair at a fixed distance from a loudspeaker. The HRTFs were then collected by rotating the chair to 12 predefined positions, and repeating this for 6 different speaker heights, arriving at 72 sampled directions in ≈ 30 min.

The dynamic measurement approaches differ on several key points with the SNAPSHOT measurement technique. (1) Whereas previously static, stop-and-go ‘snapshots’ were taken at fixed positions and movement during reception of the (300ms) stimulus signal was undesired, in a dynamic measurement setup this constraint is relaxed: the subject samples different directions by moving the head with respect to the single loudspeaker even during stimulus reception. (2) Moreover, the dynamic measurement does not necessary require a predefined regular sampling scheme, but can be carried out either in a supervised or unsupervised manner, because (3) the head movements (and consequently the varying directions of the loudspeaker with respect to the head) are now tracked in some way, e.g. by an Inertial Measurement Unit (IMU) attached to the head or by the 3D positioning system of nowadays virtual reality platforms [18]. And finally, (4) since the subject is allowed to bend his head forward, backward and sideways, hence sampling the upper and lower hemispheres, the loudspeaker can remain at a fixed position and should no longer be positioned at different heights, as was previously the case. Such an approach allows for a faster and more user-friendly HRTF measurement at home.

Obviously, the signal-to-noise ratio (SNR) of an HRTF obtained in a home environment is not as high as in a state-of-the-art measurement. But this may well be not a problem depending on the applications for which the obtained HRTF is used. Indeed, even if a dynamic HRTF would perform worse than a static HRTF, it may still be superior to a generic non-individualized HRTF, and thus be of added value.

In this paper, we investigate the feasibility of such a dynamic HRTF measurement in a home environment using a single external speaker. For this purpose, we make use of widely accessible, low-cost equipment. Moreover, the measurement is carried out in an echoic room, similar to an ordinary room in a home environment. The only exception to this low-cost approach is our use of a high resolution tracking system (Qualysis motion tracking system) to monitor the head and shoulder movements during the measurement. As the accuracy of this system is an upper limit for any low-cost motion tracking alternative, it allows us to study the feasibility of a dynamic HRTF measurement approach without limiting ourselves to a particular choice of tracking technology.

To highlight the differences that are due to the dynamic nature of measurement, we compare the obtained dynamic HRTF with a static one that was obtained by a standard HRTF measurement in an anechoic room. We show that the dynamic nature of the measurement results in an HRTF which is similar to, but in some respects essentially different from a static HRTF. The fact that the head moves with respect to the loudspeaker during the measurement introduces artefacts that are explicitly avoided in a static measurement. We argue that the main differences, though, arise from the fact that the head moves with respect to the torso of the subject, whereas in the standard HRTF measurement the head is static, always in the same (upright) position with respect to the torso. These differences are inherent to the measurement setups, their possible perceptual relevance needs further investigation.

II. METHOD

A. MEASUREMENT SETUP

The setup is very similar to the one used in a standard HRTF measurement facility: a subject is equipped with in-ear microphones and broadband stimulus signals are applied from an external loudspeaker. But whereas normally the subject should remain still and the (sound producing) loudspeaker changes position, here, only a single, fixed loudspeaker is used and the subject varies the relative direction of the loudspeaker by moving his/her head, see Fig. 1. The head and shoulder movements are tracked during the measurements, such that upon reception each stimulus can be assigned to a particular source direction with respect to the head. As the subject moves the head in all directions and rotates on the chair, directions on the sphere are sampled, and from the collected data, a dynamic HRTF can be extracted. The measurement setup, procedure and processing that was used for the data presented are described below in detail.

1) ROOM SETUP

Measurements were carried out in a room (dimensions 4m×3.5m×7m) with highly reflective tiled walls and cement floor. Two large curtains were hung, dividing the long axis of the room in two, to reduce sound reverberations (reverberation time $T_{30} = 0.42s$). The subject directly facing the loudspeaker was seated on an ordinary office swivel chair, its rotation axis located at a distance $d = 1.5$m from the loudspeaker, since we are interested in the far field HRTF [1].
The height of the swivel chair was adjusted such that the subject’s head was (approximately) at the same height \( h = 1.40 \text{m} \) as the loudspeaker (Fig. 1).

To carry out the measurement in such a reverberant room with limited dimensions, one has to deal with the reflections of the sounds bouncing of the floor, walls and ceiling. This is achieved in part by choice of an appropriate stimulus signal, see later. In addition to this, it is necessary that the measurement setup is organized such that the close surroundings of the head and the loudspeaker are free from reflectors, see Fig. 1(a). Indeed, the time between arrival of the direct path and first reflection will ultimately determine the maximal duration of the Head Related Impulse Response (HRIR) that can be measured. In general, if the distance between head and loudspeaker is \( d \) and we would like an HRIR of length \( w \), then it would require an ellipsoidal volume free of reflectors, the head and the loudspeaker in the focal points \((h, \pm d/2)\), such that for all points outside the ellipsoid, the sum of the distances to the focal points exceeds the distance \( d \) between the focal points by at least \( w \).

In our setup, the loudspeaker and head (at starting position) were at a height \( h = 1.40 \text{m} \) and a distance \( d = 1.5 \text{m} \) apart which allows for a reflection free HRIR length \( w \leq (d^2 + 4h^2)^{1/2} - d = 1.44 \text{m} \). (Note that if \( d \) increases, the possible HRIR length \( w \) decreases, while for \( h \) it is the opposite.) Since the centre of the head moves during the measurement, an HRIR length of 1.24m was assumed, which corresponds to a duration of 3.6ms and a frequency resolution of \( \Delta f = 279\text{Hz} \) (assuming a sampling frequency of 44.1kHz) for the HRIR derived from such an HRIR, which is similar to the frequency resolution used in common HRIR databases (cfr. CIPIC: 223Hz) [2]. Note that the dimensions of the object-free ellipsoid are by no means uncommon. A similar frequency resolution can therefore also be obtained in an ordinary home environment.

2) AUDIO PRODUCTION

A low-cost single driver loudspeaker (JBL Go, single driver, 40mm diameter, frequency range: 180Hz-20kHz) was used in the measurement; the stimulus sound file was encoded as a 44.1kHz wav-file and played through the loudspeaker via a Sandisk Clip Sport audio Player. As stimulus an exponential frequency-modulated sweep from \( f = 300\text{Hz} \) to 22kHz was used [19], covering the full hearing range, except for the low frequencies. Because of the dynamic nature of the method, the duration of the stimulus had to be short to minimize motion-artefacts as the subject moves the head during reception of the stimulus. On the other hand, it could not be too short, to achieve a sufficiently high SNR across the relevant frequency range. We settled upon a stimulus duration of 28.6ms.

The broadband stimuli were interleaved with periods of silence, to make sure that the sound reverberating in the room was sufficiently attenuated, see Supplementary Material.
The head movements were continued during the full 15 min while not moving his torso (and shoulders) on the chair. He could rotate the chair freely, but slowly, using his legs, freely in all directions (up-, down-, sideways), see Fig. 1(c-e).

4) ORIENTATION AND POSITION DATA COLLECTION
The orientation and position of the head, the torso and the (fixed) loudspeaker were measured using the Qualisys motion capture system, see Fig. 1(b), which ran on a dedicated computer. Reflecting markers were attached to the subject’s head (five markers on a cap) and shoulders (two markers for left and right) and to the loudspeaker (three markers). A total of six infra-red camera’s monitored these markers during the measurement, each from a different angle. Fusing the collected data, the three-dimensional trajectory of each of the markers was reconstructed with a resolution of \( \approx 1 \) mm at a rate of 100Hz. From these data, the instantaneous position and orientation of the head (and ears) with respect to the shoulders and the loudspeaker were obtained.

5) SYNCHRONIZATION OF THE DATA STREAMS
The audio production, reception and the motion capture system were all running on different platforms. To synchronize the captured data streams, every measurement was preceded by a simple calibration step: a table was hit a couple of times with a marked stick; the timings of the produced sounds recorded by the in-ear microphones could then be aligned with the moment of impact of the stick, as recorded by the Qualisys motion capture system. During calibration, the microphone was kept a distance of \( \approx 0.1 \) m from the point of impact, resulting in an accuracy of \( \approx 0.3 \) ms, which is negligible given the 100Hz frame rate of the qualisys motion capture system.

6) HEAD AND TORSO MOVEMENTS
The measurement was carried out on one subject. During the measurement, the subject moved his head with the objective to sample as much of the sphere as possible, while not moving too fast and while keeping the head centre as close as possible to the initial head position. To this end, the subject was given guidelines, which are detailed in the Supplementary Material. The subject had to sit straight and rotate and bend his head freely in all directions (up-, down-, sideways), see Fig. 1(c-e). He could rotate the chair freely, but slowly, using his legs, while not moving his torso (and shoulders) on the chair. The head movements were continued during the full 15 min duration of the measurement. Apart from these guidelines, the subject’s movements were not controlled in any way.

B. EXTRACTING ACOUSTIC INFORMATION
1) ISOLATING DIRECT PATH SIGNALS
Unlike in a standard HRTF measurement, the head is not at a fixed distance from the loudspeaker, as the subject moves the head during the measurement. As a result, the time window that contains the direct path stimulus is not known in advance and has to be identified and isolated for each new stimulus arrival. This is achieved by making use of the fact that (1) overall, the direct path contributions have a higher intensity than reflections (except when e.g. the contralateral ear is facing a wall), (2) the timing between subsequent stimuli deviates very little from the inter-stimulus time, since the head can only move a small distance between successive stimuli and (3) the timing between stimuli arriving in the left and right ear does not differ too much, as they are only \( \approx 0.15 \) m apart. Making use of these constraints allows to isolate the contributions corresponding to the direct path from those due to reflections.

2) BINAURAL SPECTRUM
For each stimulus \( k \) the arrival time \( t_k \) was then used to isolate the time window of the binaural audio signal \( a_{L/R}^k \) that contains the direct path contribution received in both the left (L) and right (R) ear. (Note that this window is approximately the duration of the exponential sweep, which is much longer than the HRIR length \( w \) discussed above). The complex spectrum is then given by

\[
H_{L/R}(k) = \frac{F(a_{L/R}^k)}{F(s)}
\]

where \( a_{L}^k \) and \( a_{R}^k \) are respectively the left and right audio signal of the \( k^{th} \) window, and \( s \) is the emitted stimulus signal which has been zero padded to have the same length as the audio signals \( a_{L/R}^k \).

The resulting spectra still contain the response characteristics of the speaker and the in-ear microphones. In order to compensate for this, the complex spectrum was first measured for the speaker-microphone combination, in absence of the subject and chair, resulting in \( H_{L/R}^{\text{system}} \) and \( H_{R}^{\text{system}} \). Note that both complex spectra were assumed direction independent (which was found to be a valid assumption; due to their small size the microphones are approximately omnidirectional). The complex spectra of the particular head/torso/loudspeaker configuration at time \( t_k \) then corresponded to

\[
H_{L/R}^k = \frac{F(a_{L/R}^k)}{F(s)H_{L/R}^{\text{system}}},
\]

or, in the time domain, the HRIR at time \( t_k \) equals

\[
h_{L/R}^k = F^{-1}(H_{L/R}^k).
\]
3) DISTANCE COMPENSATION
Due to the head movement, the distance to the loudspeaker
varies during the measurement, see Fig. 1(e). As the head
is located in the far field of the loudspeaker ($d > 1\text{m}$),
the impact of this distance variation on the spectrum can be
compensated for by multiplication of each measurement $k$
with a factor $d^k / d$, where $d^k$ is the distance between the
centre of the head and the loudspeaker upon reception, and
$d = 1.5\text{m}$ is the distance for which the calibration was per-
formed. Of course, the distance also varies during reception of
the stimulus signal, but given the limited speed of movement of
the head/ear towards/away from the loudspeaker (0.1m/s),
this can be safely neglected, see Supplementary Material.

C. INTERPOLATION
In a static HRTF measurement, where all directions are sampled
at a dense and regular grid, the resulting list of HRIRs
would suffice for 3D sound reproduction. Using a nearest
neighbour criterion, or taking a weighted sum of different
neighbours, one could produce an approximate, interpolated
HRTF for any given sound direction [20]. In the current setup,
though, this approach is not possible as it would result in
audible artefacts. One reason is that the HRTF is not sampled
on a regular grid. Consequently, some areas are more densely
sampled than others, while other areas are not sampled at all,
e.g. the bottom part of sphere. Another reason is that, as we
will see later, the HRIRs of neighbouring sampled directions
may differ considerably in a dynamic measurement. For these
reasons, we interpolate the HRTF over the full sphere using
spherical harmonics [21].

1) MINIMUM PHASE APPROXIMATION
In the past, different strategies have been proposed to inter-
polate the HRTF [22]. Here, we assume that a minimum
phase approximation of the complex spectra, augmented by a direction-dependent, frequency-independent interaural
time difference (ITD) contains all the relevant temporal cues
[23]. Kulkarni et al. [24] showed that this is an adequate
description of the HRTF phase, as long as the low-frequency
ITD is appropriately chosen, an assertion that was somewhat
nuanced later on, as it was shown that this approximation may
not be perceptually neutral for all directions [25] and results
in slightly inferior localization performance for directions
close to the interaural axis [26]. Nevertheless, as this effect
is small, this approximation is often used, e.g. [27], since it
has the advantage that one can straightforwardly interpolate
the ITD and the spectral magnitude values, as they are real
(i.e. non-complex) quantities.

a: INTERAURAL TIME DIFFERENCE
Since all relevant phase information is considered to be
encoded in the ITD, it is crucial that the latter is assessed cor-
rectly. The ITD is defined as the difference in time of arrival
of the wave front originating from a single sound source at the
listener’s ears. Unfortunately, there is no absolute definition
of the ‘time of arrival’, and a variety of methods for extracting
the ITD have been proposed, [28], [29] each having their own
interpretation of ‘time of arrival’. These different methods
lead to perceptually different results. However, it was shown
that the perceptually most relevant procedure across vari-
ous metrics is consistently the first-onset threshold detection
method. In this method, the time of arrival is obtained by
detecting the crossing of a low relative threshold on a low-
pass filtered version of the HRIR (cut-off frequency: 3kHz).

Therefore, the same approach is used here: after filtering the
measured HRIRs of the $k^{th}$ sample $h_{L/R}^k$ with a Butter-
worth low-pass filter (order 6, 3kHz cut-off), the timing of the
respective left and right onsets $t^L_k$ and $t^R_k$ are obtained by
detecting the crossing of a threshold of $-10\text{dB}$ relative to the
peak value, as used in [30]. The ITD of the $k^{th}$ sample is then
calculated as $\text{ITD}^k = t^L_k - t^R_k$.

b: SPECTRAL MAGNITUDE
Because the sensitivity of our perceptual system is logarith-
mic, in the following, the magnitude is always expressed in
dB, i.e. $20\log_{10}|H_{L/R}^k|$ [27].

2) SPHERICAL HARMONICS REPRESENTATION
The preceding steps lead to irregularly sampled ITDs and spectral magnitudes. In order to interpolate these quantities
over the full sphere, they were expressed in a truncated basis of real spherical harmonics $Y^m_l$, with $-l < m < l$
[27]. Spherical harmonics are the equivalent of Fourier basis
functions, but now defined on the full sphere. The higher the order, the more spatial detail can be accounted for. Hence,
truncation imposes an upper limit on the spatial detail that
can be captured by representations using this limited set of
basis functions.

For each of the different frequency bands with central frequency
$f$, the magnitude $H^k(f)$ corresponding with sampled direction $(\theta^k, \phi^k)$ was projected on a spherical harmonics
basis

$$\tilde{H}^k(f) = \sum_{l=0}^{L} \sum_{m=-l}^{l} C^m_l(f) Y^m_l(\theta^k, \phi^k),$$

with truncation order $L$. This entails estimating for each
frequency $f$ a set of coefficients $C^m_l(f)$ for which the sum of
the squared residual errors

$$\sum_k [\epsilon^k(f)]^2 = \sum_k [H^k(f) - \tilde{H}^k(f)]^2$$

is minimal, where the residual error $\epsilon^k(f)$ is defined as the (signed) difference between the measured magnitude
$H^k(f)$ and its spherical harmonics fit $\tilde{H}^k(f)$.

In case of a regular grid over the full sphere, it was shown
that a decomposition of an HRTF into a spherical harmonics
basis with truncation order of $L = 10$ results in a residual
error $\epsilon(f)$ which is smaller than $2\text{dB}$ for frequencies below
13kHz [27], and has no perceptual impact on spatial hear-
ing. However, because of irregular sampling of the sphere
and the fact that some parts of the sphere have not been sampled at all, regularization problems may occur, even in case of $L = 10$. These problems occur if the non-sampled areas are comparable in size to the spatial variations of the highest order spherical harmonics basis functions. To address these regularization problems, a common strategy is to use Tikhonov regularization [31] and minimize both the residual error and the norm of the coefficient vector $C_l^m$. We found that in the current case this is not a useful criterion, because if data is missing in some parts of the sphere (e.g. in the lower hemisphere, which is difficult to sample, see further), minimization of the norm of the coefficient vector will minimize the function values $H_k(f)$ in these areas. Hence, there is a bias towards ‘zero’ magnitude in those parts of the sphere that have not been sampled, whereas a smooth interpolation of the values obtained on the boundary of this region is preferred.

For this reason, in addition to the norm of the residual vector, it was opted to minimize the norm of the coefficient vector but only for values with $l > 2$, i.e., to find a solution for which

$$
\sum_k [\epsilon_k^f(f)]^2 + \lambda \sum_{l=3}^{L} \sum_{m=-l}^{l} [C_l^m(f)]^2
$$

was minimal ($\lambda$ being the regularization parameter, set to 4). The reasoning is the following: by not penalizing energy $[C_l^m(f)]^2$ that is contained within coefficients with $l \leq 2$, there is a bias towards using these lower spherical harmonics functions in the representation. These functions vary on a large spatial scale, i.e. larger than the gaps in the data and as a consequence, the function values in these non-sampled regions will no longer be pushed towards zero, but instead vary slowly.

For the interpolation of the ITD we use the same strategy (resulting in spherical harmonics fit ITD$^k$ and residual error $\epsilon_k^{ITD}$), but as the ITD varies more slowly, we consider a truncation number of $L = 5$.

### III. RESULTS AND DISCUSSION

#### A. STATIC HRTF

The static HRTF of the subject for which the dynamic HRTF was obtained, was also measured in a specialized laboratory [9]. In a semi-anechoic room, the subject’s head was centred in an vertical arc, on which 65 loudspeakers were mounted (2.5° elevation resolution between $[-70°, 90°]$). The full-sphere HRTF was obtained by rotating the platform (5° azimuth resolution) on which the subject was standing, resulting in a very dense, regular grid of 4680 sampled directions, see Fig. 3(a), where the right ear spectral magnitude is shown for $f = 8094$Hz.

#### B. DYNAMIC HRTF

1) SPHERE COVERAGE

From the different sampling directions shown in Fig. 2(b), it is clear that, given the physical constraints of the human body, it is possible to sample most of the sphere using the setup considered. The subject had normal mobility of the head/neck. If the subject were to be hampered in any way, e.g. because of neck problems, the overall coverage of the sphere would be less. Due to the unsupervised data acquisition, the distribution of sampled directions is irregular and different for every measurement. With $\approx 3000$ directions, the sphere is rather densely sampled, though some areas are more densely covered than others.

2) ITD

The measured ITD is shown for the different sampled directions in Fig. 2(b). If the ITD is expressed in spherical harmonics with truncation order $L = 5$, the residual error $\epsilon^{ITD}$ (defined as the difference between measured and fitted ITD) is shown in Fig. 2(c) and has a distribution with $\sigma = 0.019$ms, see Fig. 2(d). Overall the spherical harmonics representation fits the measured ITDs rather well. The residual errors are largest on the lateral sides at low elevations.

---

**FIGURE 2.** (a) Schematic view of the spherical projections used throughout this manuscript. The left and right hemispheres are plotted separately, using a Lambert azimuthal equal area projection method. (b) The ITD as measured for the sampled directions in the dynamic measurement approach. (c) The residual error, when the ITD is expressed in a spherical harmonics basis ($L = 5$) and (d) the distribution of the residual errors.
FIGURE 3. The right ear spectral magnitude at \( f = 8094 \text{Hz} \) for the sampled directions, (a) for the static HRTF, (b) for the dynamic HRTF as measured and (c) when projected to a spherical harmonics basis \( (L = 10) \); (d) the standard deviation of the residual error when expressed in a spherical harmonics basis, calculated for each direction over all frequencies ranging from [0.5-12kHz].

3) SPECTRAL MAGNITUDE

With every sampled direction corresponds a binaural magnitude spectrum. The right ear magnitude is shown in Fig. 3(b) for \( f = 8094 \text{Hz} \). Compared to the static HRTF shown in Fig. 3(a), the magnitude of the dynamic HRTF shows more variability, as it changes more abruptly between neighbouring sampled directions. In Fig. 3(c) the spectral magnitude is expressed in spherical harmonics basis with \( L = 10 \). A low truncation order imposes spatial smoothness on the magnitude fit, and consequently, the variability can be quantified by the residual errors \( \epsilon_H^2 \). In Fig. 3(d) the standard deviation of the residual errors of the right ear magnitude over the frequency range [0.5kHz-12kHz] is calculated for every sampled direction. The residual errors are largest (i.e. the magnitude varies most) for lower elevations, are slightly smaller on the sides and in the front, and are smallest in the back. Different effects contribute to this spectral variability. The different effects and their relative weights are studied below.

C. HEAD-TORSO CONFIGURATION

The main difference between the proposed dynamic HRTF measurement and the static measurement is that the subject moves his/her head during the measurement. This results in variations in the HRTF, see Fig. 2(c) and Fig. 3(d) that are not present in a static HRTF measurement. We argue that these differences are mostly due to the changing orientation and position of the head with respect to the torso during the measurement. Indeed, seen from the reference frame of the head, two neighbouring sampled directions can correspond to completely different head/torso configurations, as illustrated in Fig. 4. Here, we show two neighbouring sampled directions from a part of the sphere where the spectral variations (residual errors) are large, i.e. from the encircled area shown in Fig. 2(c) and Fig. 3(d). Though the selected directions differ only 0.7°, they correspond with a different torso configuration (torso azimuth angle difference is 112°) and consequently their spectral magnitudes are different, see Fig. 4(c). Overall, the magnitudes are similar, yet, as is especially apparent for the contralateral ear, a comb-like filter is superimposed due to interference with delayed reflections on the torso, see Fig. 4(d).

The effect of the torso on HRTFs was extensively studied by Algazi et al. [32] in case of the standard fixed head-torso
orientation and by Brinkmann et al. [33] in case of variable head-torso orientations. These authors distinguished between two different ways the torso affects the HRTF: shadowing and reflection. If the torso is blocking the direct path between source and ear, it acts as a shield and attenuates sounds with frequencies above 100Hz up to 25dB. In addition, the torso can also act as a reflector, creating secondary reflections that add to the direct path sounds, causing comb-like filters with amplitudes up to ±5dB. The exact positions of the peaks and notches of the comb-filter depend on the timing of the reflection with respect to the direct path, and consequently carry mainly information about the elevation of the sound source. This can also be understood by looking at the peaks and notches, i.e. the interference pattern, shown by the static HRTF in Fig. 3(a). They show the directions for which the comb-filter attains a peak or notch for this particular frequency. As can be seen, the peaks and notches are aligned along similar elevations, and hence relay information on the height of the sound source. In case of a source right above the head, the first notch already occurs at 700Hz. Hence, the influence of the torso is felt throughout the full audible range, whereas pinna cues become salient only for frequencies above 3kHz [33].

Shadowing is important for sources at lower elevations, below −40° as it can introduce attenuation up to 25dB. This explains the large spectral variations and the resulting large residual errors for lower elevations shown in Fig. 2(c) and Fig. 3(d). Indeed, the torso may have been blocking the direct path for one sampled direction, while for its neighbouring direction this was not the case, which can create large variations in neighbouring spectra. Because shadowing operates at lower frequencies (starting at 100Hz) and the ITD was calculated based on time of arrival of low-frequency HRIRs, it also affects the ITD at lower elevations most, as is clear from Fig. 2.

A similar reasoning holds in case of torso reflections. If the torso is at a different orientation the reflection may differ in amplitude and in path length. As a consequence, the pattern (position and height/depth) of the peaks and notches of the comb-filter may be different for two neighbouring directions, again resulting in spectral variations. At the frequency shown, a different torso orientation will result in a peak/notch interference pattern different from the one shown in Fig. 3(a), and consequently Fig. 3(b) should be seen as a sampling of a collection of different interference patterns. The impact of comb-filtering was shown to be strongest on the contralateral side [32], [33], which explains the larger contralateral residual errors shown in Fig. 3(d). This is likely due to the fact that for the contralateral side, torso reflections are stronger relative to the direct path sound.

It should be noted though that the size of the spectral variations were slightly larger than the ±5dB comb-filter pattern due to torso reflections documented before [32], [33]. This is due to the fact that in the dynamic HRTF measurement the head could also bend forward, backward and sideways, while in previous analyses the head was always in the upward position and only azimuthal rotations were allowed. In order to sample the sphere adequately, the subject had to move the head much closer to the torso/shoulders. As a consequence, more sound energy is being reflected, causing higher (deeper) peaks (notches) of the comb-filter, which explains why the variations exceeds ±5dB reported earlier.

**D. OTHER SOURCES OF SPECTRAL VARIATION**

In addition to the varying head-torso configuration (shown in Fig. 5(a)), there are other sources of variations, which are also due to the head movement during the measurement. Here, we discuss how these variations arise and how these
can be reduced and/or compensated for. In Fig. 5, the relevant quantities are shown as measured on the sphere. This way the presence (or lack) of spatial correlation of the variations they introduce can be discussed. Whereas spatially decorrelated variations on the ITD or spectral magnitude may be ‘averaged out’ when projected on a truncated spherical harmonics basis, spatially correlated variations may not. Consequently, spatially correlated variations may result in systematic errors in the measured HRTF.

1) DISTANCE TO LOUDSPEAKER
Since the head moves, the distance $d$ to the loudspeaker changes during the measurement (with $\sigma = 0.10\text{m}$), as shown in Fig. 5(b). The distance is spatially correlated: $d$ decreases with increasing elevation. This can be understood as follows: in order to sample the upper (lower) sphere the subject had to tilt the head forward (backwards), resulting in a smaller (larger) distance to the sound source. Note that the distance also depends on the position of the subject’s head center with respect to the rotation axis of the chair. The resulting variation of the magnitude (with estimated $\sigma = 0.57\text{dB}$) can be easily compensated for, see Sec. II-B. Consequently, distance variation contributes very little to the variations on the final spectral magnitude estimates observed in Fig. 3.

2) ARC ANGULAR SIZE
Contrary to a static HRTF measurement, where measures are taken to ensure that the subject does not move his/her head during stimulus reception [4], this is not the case in the current setup. As a consequence, each of the sampled directions in Figs. 2 and 3 are in fact arcs (instead of dots). In Fig. 5(c), the angular size of the arcs covered during stimulus reception (defined as the angle between the direction at start and end of the stimulus signal) are plotted for each sampled direction. The arc length has mean $0.76^\circ$ and $\sigma = 0.34^\circ$. In general the HRTF varies rather slowly spatially, and for this reason, an angular resolution of $5^\circ - 10^\circ$ is deemed sufficient for HRTF processing [34].

Consequently, the nonzero arc angle contributes very little to the variations in the current measurement. Note that arc length depends on the head movement speed and the subjects should move their head sufficiently slow for this conclusion to remain valid, see Supplementary Material.

3) ANGLE WITH LOUDSPEAKER AXIS
As the head moves laterally, it is no longer in the centre of the beam, along the main axis of the loudspeaker as is the case in a static measurement. This results in the radiation pattern of the loudspeaker being sampled at different positions, see Fig. 1(d). The radiation pattern of a loudspeaker is frequency-dependent, and as a result the spectrum of the stimulus can change during the measurement. In Fig. 5(d), the angle with respect to the main axis is shown for every sampled direction. The angle varies between $0$ and $12^\circ$ (mean $= 4.66^\circ$ and $\sigma = 2.66^\circ$) and is spatially correlated. Indeed, some parts of the sphere could only be sampled using specific head orientations, which necessarily involved a displacement perpendicular to the direction to the speaker, either vertically (to sample the top/bottom of the sphere) or horizontally (to sample the sides).

The amount of spectral variation that is due to lateral head movement depends on the (size of) the loudspeaker, the frequency and the exact head position with respect to the main axis. Compensating for these spectral variations (cfr. the distance variations) would be difficult, because this would require exact knowledge of the loudspeaker radiation pattern and orientation, which is difficult to obtain in a non-calibrated home environment.

Hence, to estimate the size of these magnitude variations, we measured the radiation pattern of the loudspeaker in the horizontal plane, see Fig. 6(a). Assuming axial symmetry, we can then relate each of the positions (with respect to the main axis) shown in Fig. 5(d) to an estimate of the spectral error. The resulting error distribution is shown in Fig. 6(b), in case the loudspeaker axis was horizontal, as in the measurement (solid line). The spectral error $\sigma$ is below $0.2\text{dB}$ for frequencies below $10\text{kHz}$, and increases for higher frequencies.

But since the head was on average below the loudspeaker axis, the errors would have been smaller if the loudspeaker had been tilted slightly downwards. This is illustrated by the dotted line in Fig. 6(b), which shows the spectral error in case the loudspeaker would have been pointing at the (unknown) average head position during the measurement. Clearly, this would reduce the spectral error, especially for higher frequencies.

Since these variations are fairly small, certainly compared to the ones introduced by the torso reflections, we propose not to compensate for them but to minimize them by adhering to the following guidelines in the design of the measurement setup: (1) use a small loudspeaker, (2) point (the main axis of)
the loudspeaker slightly below the head when in the upright position (ideally at the anticipated average head position during the measurement), and (3) make sure that the subject’s head movements are such that lateral movements are kept to a minimum (e.g., the subject’s head should be as close to the chair’s rotation axis as possible).

E. COMPARISON WITH STATIC HRTF

The static HRTF was obtained on a regular grid and consequently a one-to-one comparison with the dynamic HRTF was not possible. For this reason, the ITD and magnitude were first expressed in a truncated basis of spherical harmonics, and subsequently, the HRTFs were evaluated (interpolated) at 2000 uniformly distributed directions over the sphere (every direction corresponds to a solid angle of $4\pi/2000$). Directions with elevation below $-45^\circ$ were omitted. The ITDs and HRTFs were then compared on this uniformly sampled grid. To preserve the high spatial resolution of the static HRTF, a much higher truncation order was considered for the static HRTF ($L = 35$ for both HRTF and ITD), compared to the dynamic HRTF ($L = 5$ for ITD and $L = 10$ for magnitude).

The head reference frame was slightly different in the static measurement setup from that in the dynamic measurement. Since a small difference could result in large spectral differences, the static HRTF was first rotated, such that both HRTFs were properly aligned, i.e., the ITD and magnitude differences were minimized.

Before discussing the differences between the dynamic and the static HRTF, it should be noted that there are also differences between static HRTFs when measured repeatedly. Indeed, HRTF measurements by different laboratories on the same dummy (Neumann KU-100) showed rather large variations, both in ITD and spectral magnitude [30]. This was attributed to different equipment, room acoustics, microphone positioning, etc. But even when the HRTF is repeatedly measured in the same laboratory on the same subject, the resulting HRTFs also show significant spectral variation [35]. In the following, we use the variations in ITD and spectral magnitude reported in these two studies as a baseline to interpret the observed differences between the static and dynamic HRTFs.

1) INTERAURAL TIME DIFFERENCE

The ITD differences between the static and dynamic HRTF are shown in Fig. 7(a). The difference has mean $-5\mu$s and $\sigma = 22\mu$s, see Fig. 7(b) and is largest in areas where the residual errors are largest (Fig. 2(c)), i.e. on the contralateral side at low elevations. This is as expected, since the static HRTF (ITD) corresponds with one single torso configuration. As the inter-laboratory study by Andreopoulou et al. [30] reported ITD variations ranging from 30$\mu$s to 100$\mu$s, we conclude that the static and dynamic ITDs are in close agreement.

2) SPECTRAL MAGNITUDE

Because we are interested in the directional variations of the HRTF, the diffuse field response (the direction-independent component) was subtracted from the magnitude on a per frequency basis (diffuse field equalization) [4], [35].

In Fig. 8(a) the standard deviation (across all frequencies up to 12kHz) of the magnitude differences between the static and dynamic HRTF is shown over the interpolation grid. The spectral differences are largest on the contralateral side, again as expected in the area where the residual errors are largest, see Fig. 3(d).

The nature of these spectral differences is further illustrated in Fig. 9(a,b), where the magnitude is shown in the horizontal plane (zero elevation) for the static and the dynamic (right ear) HRTF. Both HRTFs are visually very similar, but there are clear differences. The detailed interference pattern present in the static HRTF on the contralateral side (caused by constructive and destructive interference of sound
waves travelling along different paths around the head) is still present in the dynamic HRTF, but it is now superimposed on a background of variations due to changing head-torso configurations, as discussed before. On the ipsilateral side, the HRTFs differ mainly below 4kHz, where the comb-like interference pattern clearly visible in the static HRTF is blurred in the dynamic one. This is even more clear if we compare both HRTFs in the midsagittal plane, see Fig. 9(c,d). Again, for lower frequencies < 4kHz, the clear comb-like interference pattern in the static HRTF is somewhat blurred in the dynamic case, due to the variable head-torso configurations.

Regarding the size of the spectral differences, the standard deviation of the spectral differences between the dynamic and the static HRTF over the sphere are plotted for each frequency in Fig. 8(b). The standard deviation is $\approx 2$dB and increases with the frequency to $\approx 3$dB. The size of these spectral variations are comparable to the intra-subject variations of repeated HRTF measurements, reported by Andreopoulou et al. [35].

**F. SENSITIVITY ANALYSES**

To investigate the sensitivity of the acquired HRTF to the exact measurement setup, we carried out five additional measurements on the same subject, all of them (except one) with a slightly different measurement setup. These repeated measurements produce HRTFs that are all very similar, even if the distance to the sound source, the stimulus period and the stimulus duration are varied, see Supplementary Material. As a result, the ‘quality’ of the HRTF (defined as the spectral/ITD similarity with the static HRTF) seems to be insensitive to the exact values of the design parameters of the measurement setup. Differences between measured HRTFs are predominantly due to the differences in sampled directions and in head/torso configurations, rather than being caused by the specifics of the measurement setup.

**IV. LIMITATIONS OF THE STUDY AND FUTURE WORK**

Measurements in this study were carried out in one reverberant room only. One could argue that, before making claims as to whether a dynamic measurement can be carried out in any home environment, one should test different home environments. Yet, we show that the precise conditions of the environment are not important, as long as the measurement room allows for the setup shown in Fig. 1. In the setup, we make sure that reflections on floor, ceiling and other objects are not taken into account. This is the case in any other room, by construction, as long as the measurement conditions outlined in Fig. 1 are met, which is possible in most homes. The only relevant property that can vary between rooms is the reverberation time, but this can be dealt with by increasing the inter-stimulus duration. For some rooms, the used 300ms.
inter-stimulus time may be too short, while for others it may be longer than necessary. Hence, before carrying out the measurement, it suffices to do a reference measurement of the setup in the room, as in Fig. 1 in the Supplementary Material, and then decide on the necessary inter-stimulus time. For rooms with longer reverberation times this would result in measurement durations >15mins.

This study presented results for a single test subject only. Yet, since the HRTF is such a highly dimensional property, the demonstrated similarity with the static HRTF cannot be attributed to coincidence. Moreover, there is no reason why the quality of the acoustic measurement, given a similar sampling of the sphere, would be different for another subject. The major difference between subjects, as already mentioned before, will lie in their ability to sample the sphere. Indeed, some subjects’ neck may be less flexible and as a result he or she may not attain a similar coverage of the sphere. It is therefore of great interest to study the variation of subjects’ ability to sample the sphere, and, subsequently, to investigate how this lack of coverage could be either minimized, e.g. through altering the height of the speaker, or compensated for, e.g. by using a priori information from an HRTF database. This is a topic of future research.

Because the proposed dynamic HRTF measurement setup is inherently less ‘controlled’ than a conventional static HRTF measurement in laboratory conditions, we decided to show feasibility of the approach first, before attempting to optimise the measurement procedure. This meant we chose sensible values, with no claim on their optimal nature, for a number of the design parameters: the type of loudspeaker, the positioning of the speaker (height and distance), the type of stimulus signal (duration and/or shape), the duration of data collection, the subject’s head movements (supervised or unsupervised, head speed). The impact of each of these design choices on the produced HRTF may of course be relevant and is worthy of further investigation, yet in this work, we limit ourselves to a feasibility study. From the results we conclude that, using the presented post-processing steps, at least with this particular setup a dynamic HRTF measurement is feasible within a 15min time span. However, there are two important qualifications to be made. First, we assume accurate head orientation information to be available and it still needs to be shown that such information can also be collected with a user-friendly and low-cost measurement system. Second, we showed that varying head-torso configurations gave rise to systematic differences between a dynamically and a statically obtained HRTF and the perceptual impact of such differences is a topic of future research.

Although the final goal of this research is to develop a low-cost HRTF measurement procedure for the layman user, it can be objected that the current feasibility study uses a sophisticated and expensive tracking system. Hence, we are currently developing a low-cost head orientation measurement system based on an inertial measurement unit (IMU), see [15] for preliminary results, that could be an alternative for the tracking system used in this study.

In this feasibility study, we have evaluated the quality or similarity of HRTFs in terms of the spectral magnitude and ITD differences. Ultimately, of course, there is only one valid way to assess the quality/similarity of an HRTF: by means of perceptual evaluation. Consequently, we are currently not in a position to draw conclusions on the perceptual relevance of the reported differences between a static and a dynamic HRTF, which we see as an important topic for future research.

This is all the more so, as we believe that the documented differences between both HRTFs do not necessarily mean that the dynamic HRTF is perceptually inferior to the static HRTF. Indeed, when the subject is allowed to move the head in a virtual audio display (VAS), the static HRTF is as much an approximation of the true HRTF, as the dynamic HRTF would be. In the case of using a static HRTF one assumes that the torso moves with the head, as if rigidly fixed to the head, whereas in the case of using a dynamic HRTF one assumes that humans are most sensitive to that part of the HRTF that remains invariant under head movements with respect to the torso. Hence, this raises the interesting question: which of the two HRTFs is the better approximation of reality?

V. CONCLUSION

In this paper we studied the feasibility of a dynamic method to measure the HRTF at low-cost in a home environment, using a single fixed loudspeaker and a head tracker. From the detailed comparison with a static HRTF obtained in a standard measurement setup, we conclude that such a dynamic HRTF is very similar but also that it shows systematic differences.

In part these differences can be considered imperfections due to the less controlled, dynamic nature of the measurement. The movement of the head during the measurement affects both the ITD and the spectral magnitude, because (1) the distance between head and source varies, (2) the head moves during stimulus reception, (3) and the head samples the radiation pattern of the loudspeaker for different directions. However, when certain design rules are taken into account, it was shown that these errors can be largely avoided. The remaining differences are then due to the impact of head-torso reflections.

These differences due to the variable head-torso configurations are fundamentally different from mere measurement errors, as they are intrinsic to the very concept of the dynamic measurement and therefore cannot be eliminated. Since both static and dynamic HRTFs are approximations, it remains an open question as to which of these two HRTFs is the better perceptual approximation of reality.

ACKNOWLEDGMENT

The authors would like to thank Erik Verreycken for his help with the Qualisys measurements.

REFERENCES

[1] J. Blauert, Spatial Hearing: The Psychophysics of Human Sound Localization. Cambridge, MA, USA: MIT Press, 1997.
[2] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, “The CIPIC HRTF database,” in Proc. IEEE Workshop Appl. Signal Process. Audio Acoust., Oct. 2001, pp. 99–102.

[3] F. Brinkmann, A. Lindau, S. Weinzierl, S. van de Par, M. Müller-Trapet, R. Opdam, and M. Vorländer, “A high resolution and full-spherical head-related transfer function database for different Head-Above-Torso orientations,” J. Audio Eng. Soc., vol. 65, no. 10, pp. 841–848, 2017.

[4] C. Armstrong, L. Thresh, D. Murphy, and G. Kearney, “A perceptual evaluation of individual and non-individual HRTFs: A case study of the SADIE II database,” Appl. Sci., vol. 8, no. 11, p. 2029, 2018.

[5] W. G. Gardner and K. D. Martin, “HRTF measurements of a KEMAR,” J. Acoust. Soc. Amer., vol. 97, no. 6, pp. 3907–3908, Jun. 1995.

[6] F. L. Wightman and D. J. Kistler, “Headphone simulation of free-field listening. I. Stimulus synthesis,” J. Acoust. Soc. Amer., vol. 85, no. 2, pp. 858–867, Feb. 1989.

[7] F. L. Wightman and D. J. Kistler, “Headphone simulation of free-field listening. II: Physophysical validation,” J. Acoust. Soc. Amer., vol. 85, no. 2, pp. 868–878, Feb. 1989.

[8] C. C. Berger, M. Gonzalez-Franco, A. Tajadura-Jiménez, D. Florencio, and Z. Zhang, “Generic HRTFs may be good enough in virtual reality. Improving source localization through cross-modal plasticity,” Frontiers Neurosci., vol. 12, pp. 1–9, Feb. 2018.

[9] J.-G. Richter, G. Behler, and J. Fels, “Evaluation of a fast HRTF measurement system,” in Audio Engineering Society Convention 140, New York, NY, USA: Audio Engineering Society, 2016.

[10] C. Guezenoc and R. Seguier, “HRTF individualization: A survey,” in Audio Engineering Society Convention 145, New York, NY, USA: Audio Engineering Society, 2018.

[11] D. N. Y. Zotkin, J. Hwang, R. Duraiswamini, and L. S. Davis, “HRTF personalization using anthropometric measurements,” in Proc. IEEE Workshop Appl. Signal Process. Audio Acoust., Oct. 2003, pp. 157–160.

[12] N. A. Gumerov, A. E. O’Donovan, R. Duraiswami, and D. N. Zotkin, “Computation of the head-related transfer function via the fast multipole accelerated boundary element method and its spherical harmonic representation,” J. Acoust. Soc. Amer., vol. 127, no. 1, pp. 370–386, Jan. 2010.

[13] J. Reijniers, H. G. P. Peremans, and B. W. M. Partoens, “Computer program and method of determining a personalized head-related transfer function and interaural time difference function,” WO Patent 2016 134 982, Apr. 9, 2019.

[14] S. Li and J. Peissig, “Fast estimation of 2D individual HRTFs with arbitrary head movements,” in Proc. 22nd Int. Conf. Digit. Signal Process. (DSP), Aug. 2017, pp. 1–5.

[15] J. Reijniers, B. Partoens, and H. Peremans, “DIY measurement of your personal HRTF at home: Low-cost, fast and validated,” in Audio Engineering Society Convention 143, New York, NY, USA: Audio Engineering Society, 2017.

[16] J. He, R. Ranjan, W.-S. Gan, N. K. Chaudhary, N. D. Hai, and R. Gupta, “Fast continuous measurement of HRTFs with unconstrained head movements for 3D audio,” J. Audio Eng. Soc., vol. 66, no. 11, pp. 884–900, 2018.

[17] D. S. Brungart, W. Nelson, R. Bolla, and R. Tannen, “Evaluation of the snapshot 3D head-related transfer functions measurement system,” Cincinnati Univ. Ohio, Cincinnati, OH, USA, Tech. Rep. nr 0704:0188, 1998.

[18] R. Gupta, R. Ranjan, J. He, and G. Woon-Seng, “Investigation of effect of VR/AR headgear on head related transfer functions for natural listening,” in Proc. Audio Eng. Soc. Conf., AES Int. Conf. Audio Virtual Augmented Reality. New York, NY, USA: Audio Engineering Society, 2018, pp. 1–10.

[19] A. Novák, L. Simon, F. Kadlec, and P. Lotton, “Nonlinear system identification using exponential swept-sine signal,” IEEE Trans. Instrum. Meas., vol. 59, no. 8, pp. 2220–2229, Aug. 2010.

[20] K. Hartung, J. Braasch, and S. J. Sterbing, “Comparison of different methods for the interpolation of head-related transfer functions,” in Proc. Audio Eng. Soc. Conf., 16th Int. Conf., Spatial Sound Reproduction. New York, NY, USA: Audio Engineering Society, 2019, pp. 319–329.

[21] M. J. Evans, J. A. S. Angus, and A. J. Lew, “Analyzing head-related transfer function measurements using surface spherical harmonics,” J. Acoust. Soc. Amer., vol. 104, no. 4, pp. 2400–2411, Oct. 1998.

[22] F. Brinkmann and S. Weinzierl, “Comparison of head-related transfer functions pre-processing techniques for spherical harmonics decom-position,” in Audio Eng. Soc. Conf., AES Int. Conf. Audio Virtual Augmented Reality. New York, NY, USA: Audio Engineering Society, 2018, pp. 1–10.
JAN STECKEL received the degree in electronic engineering from the Karel de Grote University College, Hoboken, in 2007, and the Ph.D. degree from the Active Perception Laboratory, University of Antwerp, with a dissertation titled Array processing for in-air sonar systems-drawing inspirations from biology, in 2012. 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, University of Antwerp, where he was the 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 Ph.D. In 2015, he became a Tenure Track Professor with the Constrained Systems Laboratory, University of Antwerp, where he researches sensors, sensor arrays and signal processing algorithms using an embedded, constrained systems approach.

HERBERT PEREMANS received the degree in electrical engineering and the degree in computer science from the University of Ghent (RUG), in 1986 and 1988, respectively, and the Ph.D. degree in electrical engineering from RUG, in 1994. Since 1999, he has been a Professor with the Faculty of Applied Economics, Department of Engineering Management, University of Antwerp (UA), where he is currently the Head of the Active Perception Laboratory. The work in the laboratory is oriented towards understanding the encoding/decoding of spatial cues in sound and its application in environment perception by autonomous systems. He was a recipient of the Marie Curie fellowship at the Artificial Intelligence department, University of Edinburgh, from 1996 to 1998.