CALCULATION OF HUMAN ECHOLOCATION CUES BY MEANS OF THE BOUNDARY ELEMENT METHOD

David Pelegrin Garcia[1], Bert Roozen, Christ Glorieux
Acoustics and Thermal Physics,
Dept. Physics and Astronomy, Katholieke Universiteit Leuven,
Celestijnenlaan 200D, bus 2416, BE-3001 Heverlee, Belgium
david.pelegringarcia@fys.kuleuven.be, bert.roozen@mec.kuleuven.be, christ.glorieux@fys.kuleuven.be

ABSTRACT

Some visually impaired people are able to recognize their surroundings by emitting oral sounds and listening to the sound that is reflected at objects and walls. This is known as human echolocation. The present paper reports the calculation of objective auditory cues present in human echolocation by means of the boundary element method using a spherical model of the human head in the presence of a reflecting disc at different positions. The studied frequency range is 100 Hz to 8 kHz. The results show that frequencies above 2 kHz provide information for localization of the object, whereas the lower frequency range might be used for size determination. It is also shown that stationary sound signals in echolocation can provide relevant acoustic cues, so as displacements in the proximity of a reflecting object become frequency-dependent amplitude modulations. Further calculations in a higher frequency range and with a realistic model of a human head could bring more light to the current knowledge in human echolocation,

1. INTRODUCTION

The hearing system has an unsurpassed capability for detecting, analyzing, discriminating and identify sounds, so that there is no electronic device which replicates its performance [1]. The process of hearing is the result of a complex interrelationship of physical, physiological, sensory-perceptual and cognitive aspects [2]. When we hear a sound, we identify several properties about the source: what/who is it? where is it? is it a potential threat? etc.

A particular case of interest is in self-generated sounds (e.g. fingersnaps, oral clicks). In this case, a person controls and manipulates them. The self-generated sound arrives to the ears of the person directly and also through the reflections at the boundaries of the environment. In principle, human beings are able to identify objects or walls by recognizing different qualities of the echoes arising in response to self-generated sounds, as the expertise of different visually impaired people like Daniel Kish or Juan Ruiz (from the World Access for the Blind [3], an organization to train blind people on using this method) proof. This object identification / localization method is sometimes referred to as 'human echolocation' in analogy to the system used by some animals, like bats or cetaceans.

Research on human echolocation was carried out in the mid-20th century by experimental psychologists as Cotzin and Dalenback [4], Kellogg [5] and Rice [6], with an interest on the factual human performance: how small and far an object can be detected, of which material it is made and what are the best self-generated signals. These experiments stated the importance of loudness and pitch cues in obstacle detection. However, no analysis was made on the acoustic signals or the transmission paths between the sound sources and the ears of the subjects.

Bassett and Eastmond [7] examined the objective variations in the sound field close to a reflecting wall. In this case, a loudspeaker playing back Gaussian noise was placed at more than 5 m from a large horizontal reflecting panel in an anechoic chamber and a microphone was placed at a number of points between the loudspeaker and the panel, observing an interference pattern. In addition, they reported the perceived pitch caused by the interference of direct and reflected sound at different distances from the wall. Even though the measurements of Bassett and Eastmond provide relevant qualitative information, they do not take into account the influence of the head and the ears on the resulting sound field.

Recently, Papadopoulos *et al*. [8] examined the acoustic cues in the experiments of Dufour *et al.* [9] using a dummy head measuring device. On the latter article, subjects had to detect the position of a panel that could be located towards the right, towards the left or in the center (in the latter case with no effective reflecting area, because its surface normal was perpendicular to the user-obstacle line). It was found that blind subjects performed much better in this task. The analysis of Papadopoulos *et al.* stated that the prevalent cues for obstacle discrimination were found in the frequency dependent Interaural Level Differences (ILD, defined as the differences in sound pressure level between the ears), especially in the range from 5.5 to 6.5 kHz, rather than on Interaural Time Differences (ITD, defined as the delay of the signals at the two ears).

Rojas *et al*. [10], [11] examined the properties of different signals used in echolocation. Whereas the palatal clicks were found optimal in repeatability and noise immunity as echolocation signals, hand claps were useful to examine far objects, due to their higher energy – and despite their lower repeatability and their misalignment with respect to the ears. In some situations, visually impaired people use hissing sounds as echolocation signals. The magnitude spectrum of anechoic

[1]Work funded by Research Foundation – Flanders

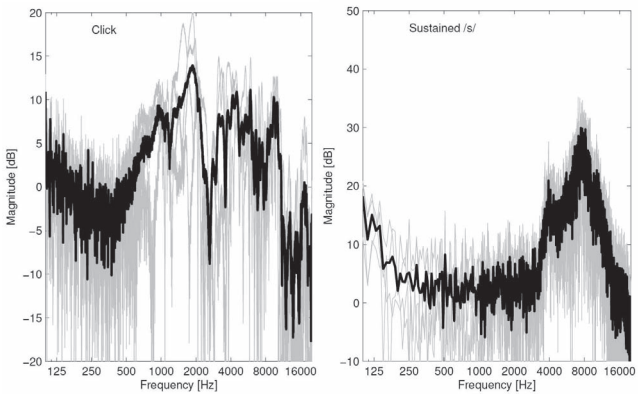

Figure 1. Magnitude spectrum of an oral click signal (left) and of a sustained /s/ sound. Individual repetitions are shown in grey and the average value is shown in black/bold.

recordings of signals is shown in Figure 1. Whereas the click signal contains energy at a broad frequency band mainly between 1 kHz and 10 kHz, the hissing sound (in the figure, a sustained /s/ sound) has energy between 4 and 11 kHz with a clearly dominant peak at around 8 kHz.

The aim of the present article is to simulate the echolocation signals at the ears – thus containing direct and reflected sound – in the presence of a reflecting object, as a complement to acoustic measurements with a dummy head. These transfer paths will be used in the future to auralize virtual objects in response to one's own voice by means of streaming convolution. The present approach looks at using the boundary element method (BEM) to calculate the transmission of sound from the mouth to the ears – and take into account diffraction effects that are oversimplified with other geometrical acoustics methods, which are more commonly used in auralization.

2. BOUNDARY ELEMENT METHOD

The boundary element method (BEM) is a numerical method used to find solutions to acoustics problems by taking into account the actual physics of sound propagation in a homogeneous medium. BEM in acoustics requires the meshing of the boundaries (but not of the medium, as finite elements methods require), the definition of boundary conditions at all points in terms of pressure, velocity or impedance, and characterization of other sound sources. BEM is widely used for solving exterior problems, namely sound radiation and diffraction, where the volume of the calculation domain is infinite.

The fundamental principle underlying BEM is the formulation of an acoustics problem as a Boundary Integral Equation (BIE). Thus, the inhomogeneous Helmholtz wave equation, states that the pressure p at position x in the presence of a monopole source with strength Q_s at position x_s varies according to

$$\nabla^2 p(x) + k^2 p(x) = -Q_s \delta(x - x_s)\,, \tag{1}$$

where k is the wave number. Applying Green's theorem to (1) within the domain V bounded by S, the direct integral formulation reads as

$$c(x)\,p(x) = -\int_S \left[\mathrm{i}kz_0 v_n(x_0) G(x_0, x) + p(x_0) G'(x_0, x) \right] dS(x_0) + Q_s G(x, x_s) \tag{2}$$

with $\mathrm{i} = \sqrt{-1}$, z_0 the characteristic medium impedance, $v_n(x_0)$ the velocity normal to the surface at the boundary point x_0. In a 3D-space, $G(x_0,x)$ is the free space Green's function defined by

$$G(x_0, x) = \frac{\mathrm{e}^{-\mathrm{i}kr}}{4\pi r} \tag{3}$$

and $r = |x - x_0|$ is the distance between two points. The factor $c(x)$ in (2) has the value ½ at smooth boundary points, 0 outside the domain and 1 at points of the domain V. At other non-smooth points, $c(x)$ fulfills the relation

$$c(x) = 1 - \frac{1}{4\pi} \int_S \frac{\partial}{\partial n} \left(\frac{1}{r} \right) dS(x)\,. \tag{4}$$

The integrals are numerically evaluated from a set of discrete points at the surface. In order for the results to be meaningful, the spacing between elements on the boundary should not be more than 1/6$^{\text{th}}$ of a wavelength. Moreover, boundary conditions of the type

$$\alpha p + \beta v_n = \gamma \tag{5}$$

(with α, β and γ arbitrary constants) must be imposed on the boundary elements. Further information about BEM can be found in [12] and [13].

3. METHOD

In this article, the sound propagation between the mouth and the ears of a spherical head model in the presence of a circular obstacle is calculated.

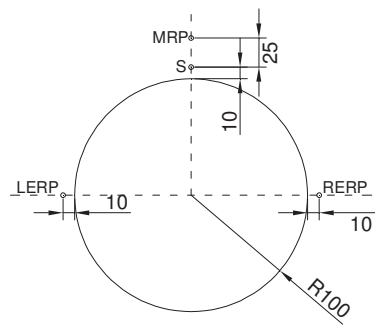

Figure 2. Model of the spherical head of 100 mm radius with the position of the source S as a mouth, and the calculation points MRP (mouth reference point), RERP (right ear reference point) and LERP (left ear reference point). Annotations are in mm.

Figure 2 shows the head modeled as a sphere of radius 10 cm; the mouth is then modeled with a monopole source at point S, 1 cm in front of the sphere, and three reference points are introduced for further calculations:

- ➔ Mouth reference point (MRP), 25 mm in front of the mouth
- ➔ Right ear reference point (RERP), 10 mm at the right of the head
- ➔ Left ear reference point (LERP), 10 mm at the left of the head

3.1. Scenarios

Three different kinds of situations — illustrated in Figure 3— were considered for calculating the Oral-Binaural Transfer Function (OBTF). In all of them, a reflecting disc of 15 mm thickness was placed in front of a totally reflecting spherical head.

In the first situation — Figure 3(a) — a disc of diameter d = 0.2, 0.4 or 0.6 m, was placed on-axis in front of the spherical head at distances between 0.2 and 2 m, in steps of 2 cm. The axis of the disc was aligned with the center of the mouth and the head.

In the second situation, a disc of 0.6 m diameter was placed in front of the spherical mouth at an on-axis distance r = 0.5 or 1 m. The OBTF was calculated for different horizontal displacements (offsets) in the direction perpendicular to the axis, in steps of 2 cm up to 1 m either to the right or to the left.

In the third situation, a disc of 0.6 m diameter was placed in front of the spherical mouth at an on-axis distance r = 0.5 or 1 m. The OBTF was calculated for different vertical displacements (offsets) in the direction perpendicular to the axis, in steps of 2 cm up to 1 m either up or down.

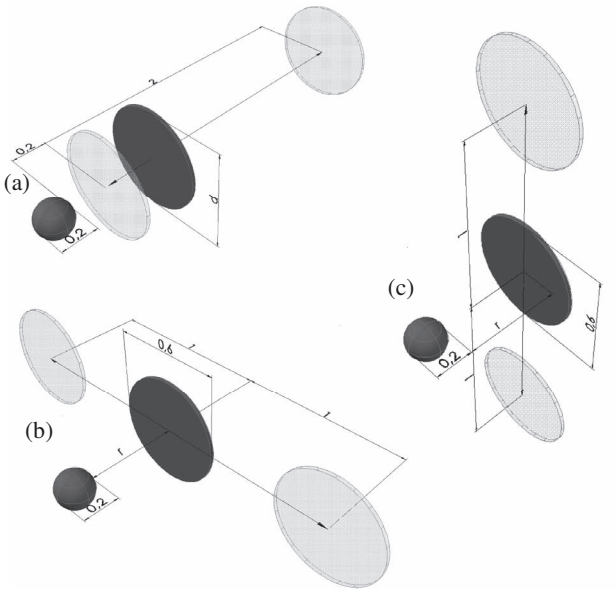

Figure 3. (a) Disc of diameter d at different distances (0.2 m to 2 m) in front of the spherical head, with its axis aligned to the mouth. (b) Disc of diameter 0.6 m at an on-axis distance r and a horizontal offset between -1 and 1 m. (c) Disc of diameter 0.6 m at an on-axis distance r and a vertical offset between -1 and 1 m.

3.2. Software

The pre-processing — i.e. meshing, adjusting boundary conditions, setting sources, field points and all other relevant parameters - was carried out with a combination of the open source project Salome [14] and Matlab [15] scripts. All the normals were pointing inwards the objects, so as to calculate the solution to the exterior problem. The sphere and the discs were modeled as meshes of triangular elements with 8 to 10 mm of spacing between neighbor nodes.

The BEM calculations were carried out with FastBEM Acoustics [16], which is a commercial BEM software that implements the Fast Multipole Method to speed up the calculations. We used a free license which was limited to a mesh with a maximum of 20.000 elements. In order to avoid singularities at the natural frequencies occurring of the interior of the head and the reflecting object, the HBIE (hypersingular boundary integral equation) formulation of the problem was used. We calculated the frequency responses with a resolution of $1/5^{th}$ octave up to 8 kHz, due to the limitation in the maximum number of elements.

The post-processing (visualization of the results) was done with custom Matlab scripts. Cubic interpolation was used to visualize the different sets of transfer functions as images. Nevertheless, FastBEM has a simple post-processor to visualize pressure and velocity on the mesh and on the field points.

3.3. Definition of metrics

The OBTF characterizes the propagation of sound between the mouth and the ears; therefore two transfer functions are calculated: the transmission between the mouth and the right ear ($H1$), and the transmission between the mouth and the left ear ($H2$), such that OBTF = $\{H1, H2\}$.

More specifically, $H1$ is the ratio of the pressure at RERP in the presence of an obstacle, p_i, to the pressure at MRP in free-field (without reflecting disc), p_{ane}. The sub-index i indicates the presence of a disc.

$$H1_i = \frac{p_i(\text{RERP})}{p_{ane}(\text{MRP})} \tag{6}$$

and analogously, $H2$ is the ratio of the pressure at LERP in the presence of an obstacle, p_i, to the pressure at MRP in free-field (without reflecting disc), p_{ane}.

$$H2_i = \frac{p_i(\text{LERP})}{p_{ane}(\text{MRP})} \tag{7}$$

In the results section, H1 is not shown and H2 will be referred to as the left channel of the OBTF or simply the OBTF (despite being non-rigorous).

The voice support ST_V, used as an objective measure in classroom acoustics (e.g. [17]), is an important metric to focus on the reflected sound. It is defined as ten times the logarithmic

ratio between the reflected sound energy and the direct sound energy at the ears, or as measurable variables:

$$\mathrm{ST}_{V,L} = 10\log\frac{\left|p_i(\mathrm{LERP}) - p_{ane}(\mathrm{LERP})\right|^2}{\left|p_{ane}(\mathrm{LERP})\right|^2} \quad (8)$$

$$\mathrm{ST}_{V,R} = 10\log\frac{\left|p_i(\mathrm{RERP}) - p_{ane}(\mathrm{RERP})\right|^2}{\left|p_{ane}(\mathrm{RERP})\right|^2} \quad (9)$$

where the differences in the numerators are the reflected pressures. The main limitation of ST_V is that it does not illustrate the interference between the reflected and the direct sound.

In addition, the Interaural Level Difference (ILD) is defined as

$$ILD = 10\log\left|\frac{H2}{H1}\right|^2 . \quad (10)$$

4. RESULTS

The left-ear OBTF for a circular obstacle at different distances from 0.2 m to 2 m in front of the head, as a function of frequency, is shown in the leftmost column of Figure 4. The interference patterns between the direct and the reflected sound

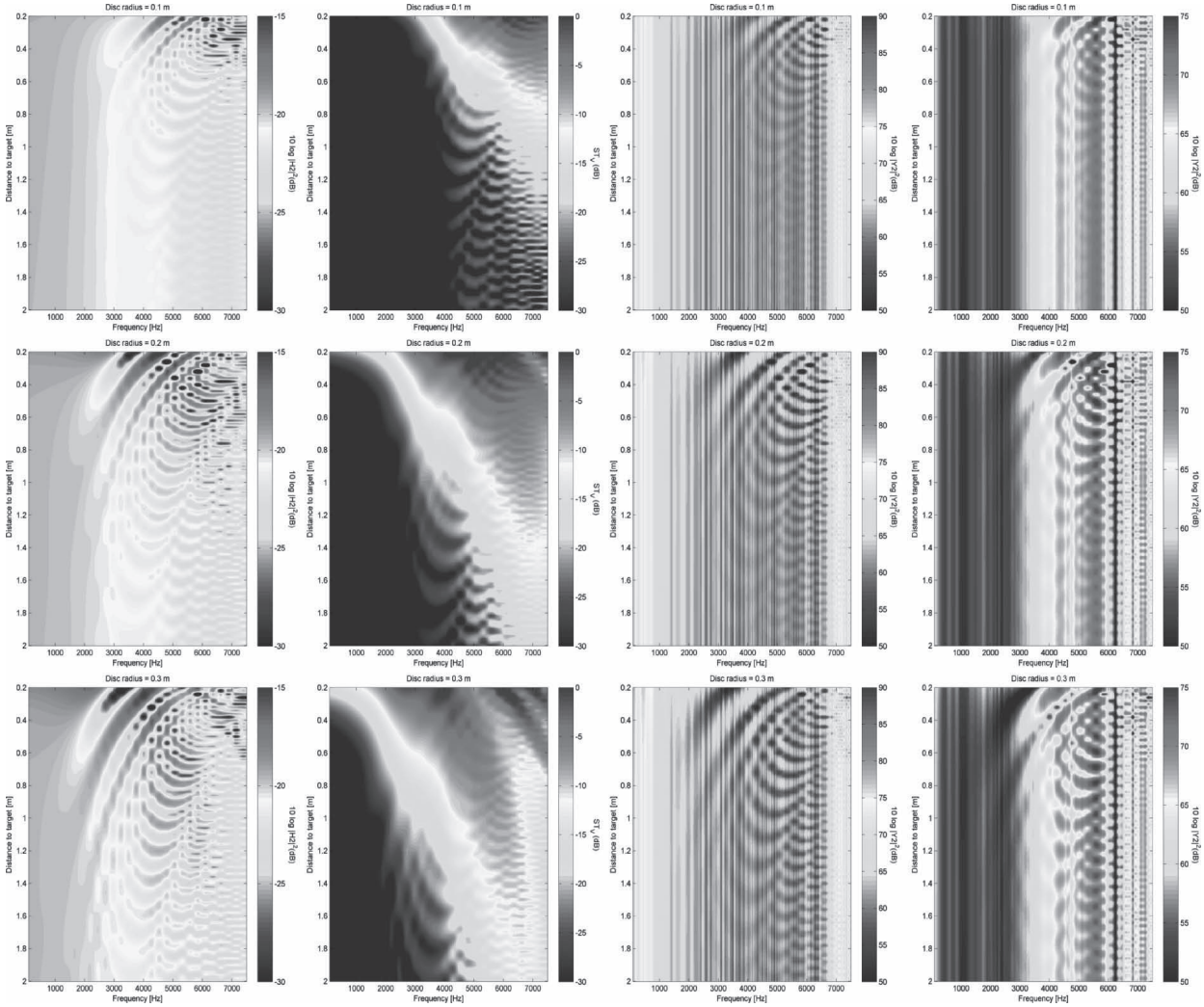

Figure 4. First column: Transfer function between MRP (free-field) and LERP (with reflector) as a function of frequency and distance, showing characteristic reflection patterns. Second column: Ratio between the energy of the reflected sound and the energy of the direct sound as a function of frequency and distance. Third column: spectrum of an oral click at the left ear with a reflector at different distances. Fourth column: spectrum of a sustained /s/ sound at the left ear with a reflector at different distances. Top row: disc r = 10 cm. Middle row: disc r = 20 cm. Bottom row: disc r = 30 cm.

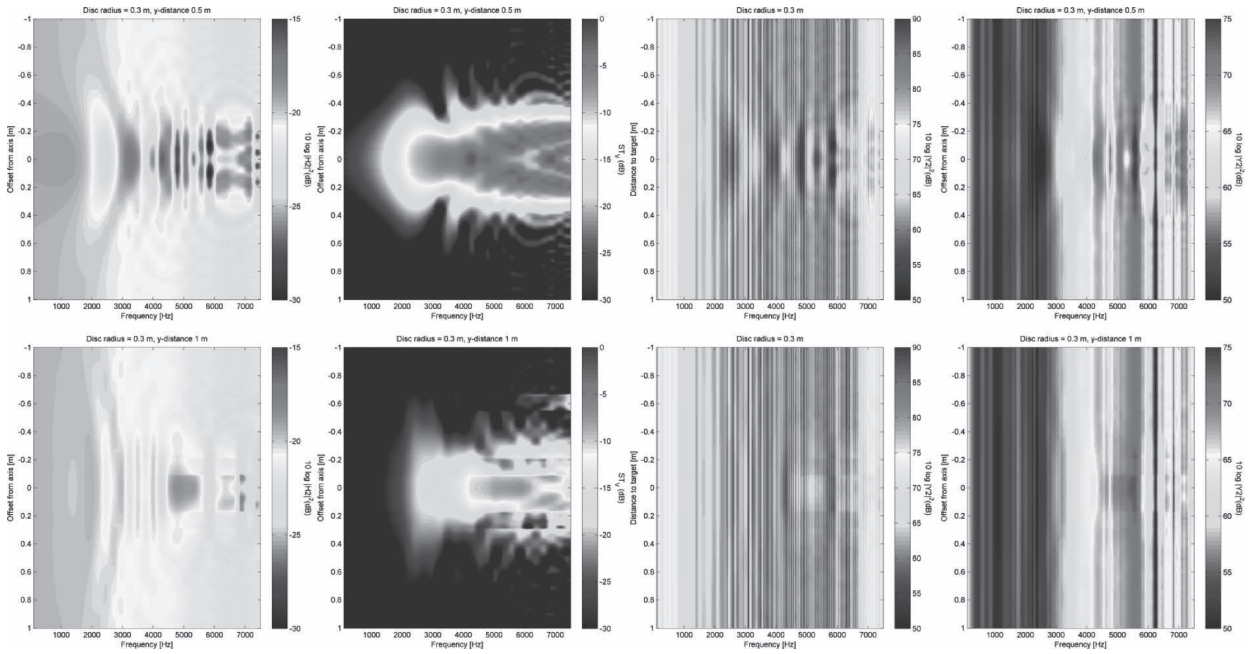

Figure 6. First column: Transfer function between MRP (free-field) and LERP (with reflector) as a function of frequency and vertical displacement of a disc of r = 30 cm. Second column: Ratio between the energy of the reflected sound and the energy of the direct sound as a function of frequency and vertical displacement of a disc of r = 30 cm. Third column: spectrum of an oral click at the left ear with a disc of r = 30 positioned with an off-axis vertical displacement. Fourth column: spectrum of a sustained /s/ sound at the left ear with a disc of r = 30 positioned with an off-axis vertical displacement. Top row: on-axis distance to the disc = 50 cm. Bottom row: on-axis distance to the disc = 100 cm.

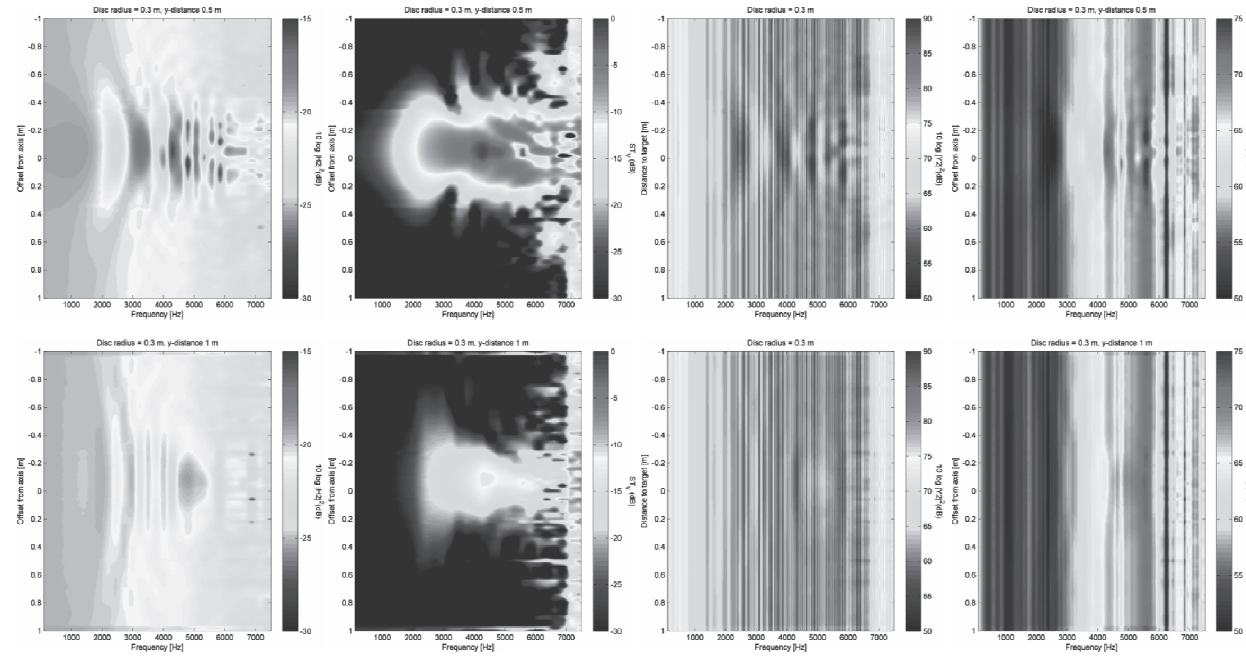

Figure 5. First column: Transfer function between MRP (free-field) and LERP (with reflector) as a function of frequency and horizontal displacement of a disc of r = 30 cm. Second column: Ratio between the energy of the reflected sound and the energy of the direct sound as a function of frequency and horizontal displacement of a disc of r = 30 cm. Third column: spectrum of an oral click at the left ear with a disc of r = 30 positioned with an off-axis horizontal displacement. Fourth column: spectrum of a sustained /s/ sound at the left ear with a disc of r = 30 positioned with an off-axis horizontal displacement. Top row: on-axis distance to the disc = 50 cm. Bottom row: on-axis distance to the disc = 100 cm.

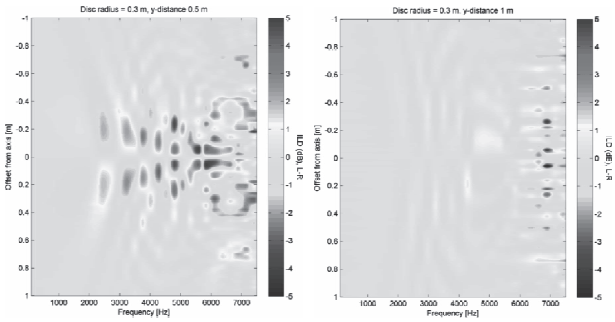

Figure 7. ILD of one's own voice in the presence of a reflecting disc (r = 30 cm) at different off-axis horizontal displacements. Left: on-axis distance 50 cm. Right: on-axis distance 100 cm.

are weakest for the smallest obstacle size (r = 10 cm, top row), whereas they become clearer with increasing obstacle size (r = 20 cm, middle row, and r = 30 cm, bottom row). The strength of the reflected sound relative to the direct sound can be observed on the second column of Figure 4. Due to diffraction effects, smaller reflectors are less efficient in reflecting the low frequencies than larger ones (compare e.g. top row, r = 10 cm, with bottom row, r = 30 cm). The actual interference patterns at the ears, taking typical echolocation signals into account is shown in the two last columns of Figure 4. The third column is obtained with an oral click as the sound source and the fourth column, with a sustained /s/ sound. It is shown that one can 'highlight' a broader spectrum of the OBTF by using an oral click rather than by using a sustained /s/.

Similar patterns are shown in Figure 5, but in this case a disc of r = 30 cm is displaced off-axis horizontally between 1 m to the left and 1 m to the right, keeping the on-axis distance constant. The distance between the mouth and the obstacle is set at 50 cm (top row) and 1 m (bottom row). Fluctuations in the OBTF are observed mainly when the obstacle is passing in front of the head. It is in the same conditions that the reflected sound is maximum (see second column, STv). The high values of reflected sound above 7 kHz indicate computational error in BEM that can be solved with the use of a finer mesh.

Figure 6 is almost identical to the previous one, showing the results for vertical displacements of a reflecting disc located at 50 cm (top row) and 100 cm (bottom row) on-axis in front of the mouth. Due to the simple geometry used in the calculations, the results are almost identical to those obtained with horizontal displacements.

Both in Figure 5 and in Figure 6 the fluctuations of the sound at the ears are very faint, indicating that discrimination across different horizontal and vertical displacements is a difficult task. However, in the case of horizontal displacements of the obstacle, there are binaural cues due to the differences between the signals at the two ears. Figure 7 shows the ILD for the different horizontal displacements of the disc, at an on-axis distance of 50 cm (left) and 100 cm (right). These cues complement the OBTF shown on the first column in Figure 5 and enable a more accurate localization. In the case of vertical displacements or on-axis displacements, there are no ILD cues, as the signal reaching both ears are identical.

5. DISCUSSION

The boundary element method is a helpful tool for the study of human echolocation, through the calculation of transfer functions of sound between the mouth and the ears (OBTF). It has been previously used for the calculation of Head Related Transfer Functions (HRTF) to external sounds, and now it is used to characterize the sound of one's own voice.

The interference patterns of the OBTF observed in Figures 4 to 6 must be judged together with the spectrum of the signals — the oral click and the sustained /s/ —, which are displayed in the third and the fourth columns on these figures. The oral click (on the third column) excites a broader range of the spectrum, so that more frequency components are available to judge the presence / absence of an object, when compared to the sustained /s/, which excites a relatively narrow band of frequencies. Nevertheless, the sustained /s/ (fourth column) is helpful in exploratory movements, because the variations in distance to the object are perceived as amplitude modulations at certain frequency bands.

Moreover, the low frequency range (below 2 kHz) does not provide any useful cues for object detection — even though the absence of energy in this frequency range might be a cue in itself for size identification. The reason for this is twofold. First, the finite object size does not reflect low frequencies efficiently — a big wall would reflect much more low frequency sound. Second, high frequency sound propagating from the mouth and the ears is shadowed by the head, whereas low frequency sound diffracts with minimum attenuation.

These figures allow estimating roughly human echolocation limits, comparing the spectral fluctuations in the figures with the outcome of amplitude modulation detection thresholds. Nonetheless, there are more acoustic cues than the ones visualized here. For example, a simple reflection of sound generates a comb filtering. In this condition, a broadband noise signal acquires a perceptible pitch sensation which is linked to periodicity of spectral notches rather than tonal components.

Figures 4 to 6 show in most of the cases smooth variations and clear patterns, which are the result of using a circular model of a head. Further variations are expected when using a realistic head model, due to the characteristic geometry of the external ear. In this case, the fluctuations of sound field, OBTFs and STv in Figures 5 and 6 and of ILD in Figure 7 are expected to be larger than shown here, and the localization of off-axis obstacles improved.

The Just Noticeable Differences in the level of fluctuating signals are of the order of 1 dB [18]. This information is useful to estimate the obstacle detectability by its comparison to the fluctuations in Figures 4 to 6. For example, a disc of r = 10 cm might not be detected at a distance of 1 m (see in the top left of Figure 4), because neither the fluctuations of the OBTF with frequency nor with distance are higher than 1 dB.

Moreover, background noise and reverberant reflections present in actual environments reduce the modulation of the OBTFs shown in Figures 4 to 6, to which echolocators would react emitting louder sounds and increasing the signal-to-noise ratio (though not under purely reverberant conditions). At some levels of background noise, due to human limitations on the

production sound and on the social acceptance of emitting echolocation signals, echolocation becomes impossible.

The frequency limit of the calculations to 7.5 kHz is representative of most of the energy content of oral clicks used as echolocation signals, as illustrated in Figure 1. Nevertheless, this limit overlooks higher frequencies which are still relevant in echolocation and where other signals contain most of their energy.

5.1. Future work

Further analysis will compare the quantities obtained with the spherical head model and those obtained with an accurate model of the head, in order to evaluate the role of the external ear in the enhancement of echolocation cues. Moreover, psychoacoustic models will be applied to the results obtained here in order to determine the relative importance of different cues in human echolocation. In addition, we will explore a higher frequency range by increasing the level of detail of the meshes used in the calculations. The data presented in this paper will be compared to the results of subjective tests for determining detection thresholds in human echolocation and under different noise and reverberation conditions.

6. CONCLUDING REMARKS

The results show the potential of BEM to simulate and predict the acoustic cues present in human echolocation in the proximity of obstacles. A simple geometrical model of the human head was used to compute the airborne propagation of sound between the mouth and the ears of the same person, in the presence of a reflecting obstacle.

It is shown that stationary sound signals in echolocation can provide relevant acoustic cues, so as displacements in the proximity of a reflecting object become frequency-dependent amplitude modulations.

7. REFERENCES

[1] R. Masterton, “Role of the central auditory system hearing: The new direction,” *Tins*, vol. 15, no. 8, pp. 280–284, 1992.

[2] J. Neuhoff, *Ecological Psychoacoustics*. San Diego, CA: Elsevier Academic Press, 2004.

[3] World Access for the Blind, “World Access for the Blind.” [Online]. Available: http://www.worldaccessfortheblind.org/. [Accessed: 14-Jan-2013].

[4] M. Cotzin and K. Dallenback, “Facial vision: The role of pitch and loudness in the perception of obstacles by the blind,” *The American Journal of Psychology*, vol. 63, pp. 485–515, 1950.

[5] W. N. Kellogg, “Sonar system of the blind.,” *Science*, vol. 137, no. 3528, pp. 399–404, 1962.

[6] C. E. Rice, “Human echo perception.,” *Science*, vol. 155, no. 763, pp. 656–664, 1967.

[7] I. Bassett and E. Eastmond, “Echolocation: Measurement of pitch versus distance for sounds reflected from a flat surface,” *The Journal of the Acoustical Society of America*, vol. 36, no. 5, pp. 911–916, 1964.

[8] T. Papadopoulos, D. S. Edwards, D. Rowan, and R. Allen, “Identification of auditory cues utilized in human echolocation—Objective measurement results,” *Biomedical Signal Processing And Control*, vol. 6, no. 3, pp. 280–290, Jul. 2011.

[9] A. Dufour, O. Després, and V. Candas, “Enhanced sensitivity to echo cues in blind subjects.,” Springer, Sep. 2005.

[10] J. A. M. Rojas, J. A. Hermosilla, R. S. Montero, and P. L. L. Espí, “Physical Analysis of Several Organic Signals for Human Echolocation: Oral Vacuum Pulses,” *Acta Acustica united with Acustica*, vol. 95, no. 2, pp. 325–330, Mar. 2009.

[11] J. A. M. Rojas, J. A. Hermosilla, R. S. Montero, and P. L. L. Espí, “Physical Analysis of Several Organic Signals for Human Echolocation: Hand and Finger Produced Pulses,” *Acta Acustica united with Acustica*, vol. 96, no. 6, pp. 1069–1077, Nov. 2010.

[12] A. F. Seybert and T. W. Wu, “Acoustic Modeling: Boundary Element Methods,” in *Handbook of Acoustics*, M. Crocker, Ed. International: Wiley, 1998, pp. 157–167.

[13] S. M. Kirkup, *The Boundary Element Method in Acoustics*. 2007, p. 148.

[14] “Salome Platform.” [Online]. Available: http://www.salome-platform.org/. [Accessed: 14-Mar-2013].

[15] “MATLAB - The Language of Technical Computing - MathWorks Benelux.” [Online]. Available: http://www.mathworks.nl/products/matlab/. [Accessed: 14-Mar-2013].

[16] Advanced CAE Research, “FastBEM Acoustics 3.0 User Guide.” Cincinnati, Ohio, US, p. 41, 2012.

[17] D. Pelegrín-García, J. Brunskog, V. Lyberg-Åhlander, and A. Löfqvist, “Measurement and prediction of voice support and room gain in school classrooms,” *The Journal of the Acoustical Society of America*, vol. 131, no. 1, pp. 194–204, 2012.

[18] J. Blauert, *Spatial Hearing*, Revised ed. Cambridge, MA and London, UK: The MIT Press, 1997, p. 494.