

# Underwater Sound Spatialization Model

Mario Michan

Electrical and Computer Engineering

University of British Columbia

2366 Main Mall

Vancouver, BC. V6T 1Z4

mmichan @gmail.com

## ABSTRACT

A technique to spatialize underwater sound has been investigated. The technique relies on bone conduction as the mechanism for sound propagation from the source to the inner ears or Cochleas. A computational physics model was developed to extract the transfer function experienced by the sound waves propagating through the skull bones. By proper manipulation of the waveforms, according to the calculated transfer function, information about the simulated position of the sound source can be conveyed to the user. Additionally, an experimental technique to evaluate the model in the lab was developed and correlated with real experiments performed underwater. The experimental technique was shown to be valid but the results show no substantial improvement on the user ability to perceive the sound position under water by using this model over the HRTF used as reference.

## Categories and Subject Descriptors

?? [Psychoacoustics]: No ACM Computing Classification Scheme Found.

## General Terms

Algorithms, Experimentation, Human Factors.

## Keywords

Human Related Transfer Function, Psychoacoustics, Sound Spatialization, Human Related Impulse Response, Underwater Sound, Echolocation.

## 1. INTRODUCTION

Echolocation or biosonar is a technique used by various animals such as bats and dolphins to locate objects in space. By emitting bursts of sound and listening to the echoes, animals can identify and range objects. Echolocation has also been observed on blind humans [1] that with the help of clicks produced by their mouths or sound produced by hitting their canes can locate large objects.

Moreover, echolocation devices based on audification of ultrasound have been demonstrated to allow operators to detect obstacles as small as 1mm. Some of these techniques take advantage of humans' capacity of sound source localization and ranging. This capacity, however, has reportedly not been developed for underwater environments. The hearing threshold and the ability to localize sound sources are considerably reduced underwater. Without this capacity humans rely on their visual and somatosensory systems to operate in this environment. Commercial divers, for example, usually operate in low visibility environments and rely on haptic techniques to locate obstacles and tools. They obtain their initial license after an average of 200 hours of training and schools usually allocate at least half of this time developing these haptic techniques.

These divers usually depend on surface observers to notify them of the direction of any approaching object. An echolocation device that can help scuba divers to operate in low visibility environments without the sole reliance on haptic perception would increase their effectiveness. A long term goal for this research is to develop such a device. However, the immediate objective of this project is to explore the fundamental theory on which such device would be based and not its full implementation. The main question that this project attempts to answer is if sound spatialization can be achieved underwater by manipulation of the sound waves.

## 2. UNDERWATER SOUND

Sounds are mechanical vibrations transmitted through a medium. These vibrations travel as longitudinal waves and transverse waves. Longitudinal or compression waves occur in gases and liquids. Acoustics describes these waves by means of physical wave properties such as frequency, wavelength, intensity, direction, etc. Psychoacoustics, on the other hand, describes sound in terms of perceptual dimensions such as pitch, loudness or timbre. Psychoacoustics also attempts to map the acoustic dimensions with the perceptual ones. Sound localization is a neural process that falls in the realm of psychoacoustics. It is, however, described in acoustical terms as achieved by mapping the interaural time difference (ITD) and the interaural intensity difference (IID) to a direction in space. This description is overly simplified since the mapping is not one to one and can result in a cone of confusion. But humans do not usually get confused. They learn at very early age how sound is filtered by its reflections and diffraction from the head, pinna and torso. This filtration is approximately described by the Head Related Transfer Function (HRTF). This treatment assumes the process is linear time

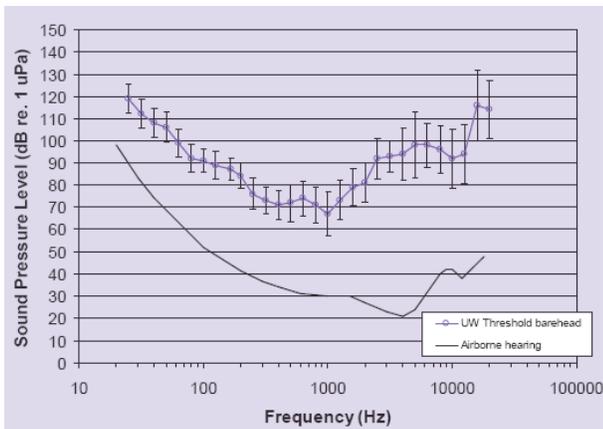
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists requires specific permission and/or a fee.

HIT2009, Vancouver, BC, Canada.

©UBC2009.

invariant and the HRTF is usually calculated by measuring the Head Related Impulse Response (HRIR) at the eardrum.

This theory is well developed and several databases with HRTF have been published. It is possible to spatialize sound and simulate a source location by processing the sound impulses with the HRTFs. In theory the same technique should be possible to use underwater. One problem is that the hearing threshold is reduced underwater. Figure 1 indicates the attenuation of the sound pressure levels at different frequencies [2]. Empirically, divers can detect frequencies between 250Hz and 6000Hz [7]. The reasons are that the resonance frequency of the external ear is lowered when the external ear canal is filled with water, and the impedance-matching ability of the middle ear is significantly reduced due to elevation of the ambient pressure the water-mass load on the tympanic membrane, and the addition of a fluid-air interface during submersion. As a result underwater hearing is mostly done by bone conduction and conduction through the ear canal is only useful for sounds below 1000Hz [3]. Another problem is that during submersion, the ITD and IID are largely lost due to the increase in underwater sound velocity and cancellation of the head's acoustic shadow effect because of the similarity between the impedance of the skull and the surrounding water. The sound velocity in air at sea level is approximately 343 m/s and in fresh water is approximately 1482 m/s.



**Figure 1: Underwater Human Hearing Thresholds**

As a result of all these constrains humans are not capable of localize a sound source underwater and a useful underwater HRTF can not be directly measured. The constrains also indicate that the most efficient way to transmit sound from an electronic device to a submerged human is through a bone conduction actuator that directly attaches to the skull.

### 3. DEFINITION

A device to enhance the human capacity to localize sound in an underwater environment will consist of two main units. The first unit will be a sound source localization device. This electronic device will consist of several hydrophones placed in different locations, filters, amplifiers, analog to digital converters, detectors and a DSP processor. The position of the sound source can be computed by comparing the time delay of the signals. Algorithms to implement this calculation based on cross correlation functions have been implemented [4]. The focus of this project is not to implement such device since it is a standard engineering problem and solutions already exist.

The second unit will be in charge of conveying the sound and its position to the operator via an auditory interface. This unit would receive the position information and the sound data from the sound source localization device and proceed to transmit it to the user. Since the interface is only auditory the sound transmitted should contain all the source sound information and its position. In order to accomplish this, the original sound waveform needs to be manipulated according to the position information. The main focus of this project is to experiment if the sound waves can be manipulated in such form that they appear to the user (which is underwater) to come from the intended location and to explore into what is involved in such manipulation. Since the application is intended in an underwater environment the transmission of sound should be via bone conduction.

To find the sound pressure that an arbitrary source  $x(t)$  produces at the ear drum, all we need is the impulse response  $h(t)$  from the source to the ear drum. This is called the Head-Related Impulse Response (HRIR), and its Fourier transform  $H(f)$  is called the Head Related Transfer Function (HRTF). The HRTF captures all of the physical cues to source localization. Once you know the HRTF for the left ear and the right ear, you can synthesize accurate binaural signals from a monaural source. Since databases of HRTF in atmosphere are readily available this is a good starting point. In order to spatialize the sound being transmitted by the underwater sound source localization device (see above), this source should be processed by the HRTF at the given position parameters. The resulting sound wave  $Y(f) = HRTF * X(f)$  will then convey the proper information of position if transmitted through airborne headphones. In this case, however, this sound is transmitted to the cochlea by bone conduction and will, therefore, be modified by the propagation mechanism. The resulting waveform will most likely be different from the intended  $Y(f)$  conveying the wrong information. We could, however, modify the waveform before it is transmitted so the resulting waveform after propagation through the bones is the intended  $Y(f)$ . Assuming that the propagation from the actuator to the Cochlea can be modeled with a linear time invariant (LTI) model (similar to the HRIR) it should be possible to calculate a Skull Transfer Function (STF). Additionally, given the LTI assumption it should also be possible to calculate an Inverse Skull Transfer Function (ISTF) such that  $STF(f) * ISTF(f) = I$ . Now, if a sound signal is convoluted with the HRTF and ISTF and then transmitted through bone conduction, then this signal will be modified during propagation by the STF and reach the Cochlea. The output at the cochlea should be the same as if the (air) HRTF was applied to the source in a normal air environment and the sound traveled through the ear canal. This manipulation inverts any modification that happens to the sound waves in the skull, which is given by the STF.

$$Input(f, \theta, \phi) \cdot HRTF(f, \theta, \phi) \cdot ISTF(f) \cdot STF(f) = Output(f, \theta, \phi)$$

The only remaining problem is to find a suitable ISTF. A straight approach would be to measure the output at the drum, which in theory should be the same as the Cochlea, when bone conduction actuators are used. This will give us the STF and from there we could calculate the inverse function ISTF. However, such measurements are very complex so a more realistic approach is to estimate a STF by simulation. For this project the STF was estimated by computational physics. A CAD model of a simplified skull was created and imported into Physics simulation software. The model was excited by applying sound waves of

different frequencies at the positions of the bone conduction actuators. The output at the Cochlea was measured in the simulation and all the parameters calculated from it. The model was created using the software Comsol Multiphysics. This software solves differential equations numerically using Finite Element Analysis. The package can solve multiple physics iteratively, which is necessary for this problem given the different mediums and densities involved in the wave propagation.

### 3.1 Model

The model uses a previously tested HRTF modified with an ISTF so sound transmitted via bone conduction could give the impression of localization. A suitable HRTF was obtained from the Listen Database [5]. The ISTF was calculated with MatLab from the STF. The STF was calculated from a numerical model implemented using the Comsol Multiphysics software. The software solves the differential wave equations of sound with proper boundary conditions on the model of the skull using Finite Difference Analysis; the source is placed in the cheekbone and the monitor in the location of the Cochlea. A Solidworks model of the skull was created and simplified to reduce computational requirements.

### 3.2 HRTF

The listen database has measurements for 51 individuals that span a range of anatomies. The HRIR measurements were performed in an anechoic room ( $8.1 \times 6.2 \times 6.45 = 324 \text{ m}^3$ ); the room is covered with 1.1m long glass wool wedges absorbing sound waves above 75Hz. Measurement equipment lay on a configurable metallic duckboard. The measurement uses a 8192 points logarithmic sweep as measurement signal (44100 Hz), and two channels inputs (one for left ear and another for right ear). The HRIR data is presented as a set of coefficients for the frequency swap and for each azimuth and elevation angle. In this case the distance is fixed.

The setup consists of 10 elevation angles starting at  $-45^\circ$  ending at  $+90^\circ$  in  $15^\circ$  steps vertical resolution. The steps per rotation vary from 24 to only 1 ( $90^\circ$  elevation). Measurement points are always located at the  $15^\circ$  grid, but with increasing elevation only every second or fourth measurement point is taken into account. As a whole, there are 187 measurement points, hence 187 stereo audio files.

Some of the published tests on this database indicate that without personalization of the HRIR only half of the users were able to externalize the sound. The perception is that of sound moving inside their heads. Personalized HRIR for 15 volunteers resulted on better performance: 8 were able to externalize and localize the source, 4 were unable to localize precisely but had success externalizing, 2 had success externalizing after training with visual feedback and 1 was unable to externalize [8]. The previous data should be taken into account when evaluating the performance of Skull Transfer Function and underwater sound localization enhancement device.

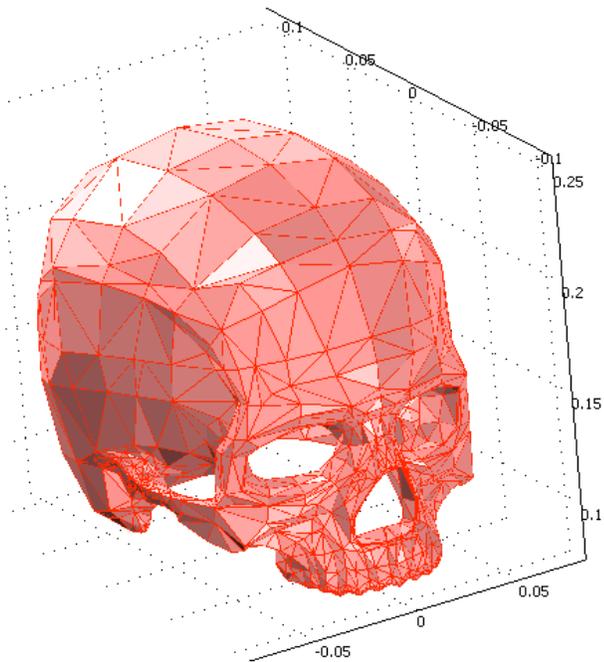
The following table shows HRIR measurement points:

**Table 1. HRIR measurement points**

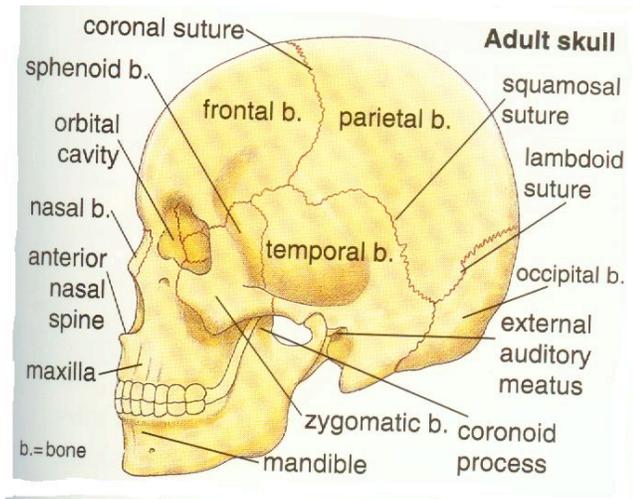
Elevation (degrees)	Azimuth increment (degrees)	Points elevation per
-45	15	24
-30	15	24
-15	15	24
0	15	24
15	15	24
30	15	24
45	15	24
60	30	12
75	60	6
90	360	1

### 3.3 Inverse Skull Transfer Function

The Inverse Skull Transfer Function is mean to cancel any modifications that happen to the sound transmitted through the bones to the cochlea. To calculate the ISTF we first calculated the STF by simulating the acoustic wave-stress strain structure interaction of the sound waves produced by the bone conduction actuators and the head and skull. Pressure waves are generated from a point source on the water-filled space next to the cheek bone (Zygomatic bones) of the skull. The sound propagates through the skull structure, the inner head and the surrounding water. See the model of the skull before simplification below and a diagram of the skull bones.



**Figure 2: CAD Model of Skull before preparation for simulation**



**Figure 3: Human Skull**

The model has two sub-domains water and solid (bone). For the water sub-domain for harmonic sound waves we use the frequency-domain Helmholtz equation for sound pressure:

$$\nabla \cdot \left( -\frac{1}{\rho_0} (\nabla p - \mathbf{q}) \right) - \frac{\omega^2 p}{\rho_0 c_s^2} = 0$$

Here, the acoustic pressure is a harmonic quantity,  $p = p_0 e^{i\omega t}$  (N m<sup>-2</sup>),  $\rho_0$  is the density (kg/m<sup>3</sup>),  $\mathbf{q}$  is an optional dipole source (N/m<sup>3</sup>),  $\omega$  is the angular frequency (rad/s), and  $c_s$  is the speed of sound (m/s). In the present model, no dipole source is included. For the solid sub-domain we calculate the harmonic stresses and strains inside skull walls using a frequency response analysis For

the bone density we use the normal human bone density for two layered bones of 1.26 (g/cm). The following is table of densities measured and references:

**Table 2. Bone density references**

Bibliographic Entry	Standardized Result
Cameron, John R.; James G. Skofronick & Roderick M. Grant. <i>Physics of the Body. Second Edition.</i> Madison, WI: Medical Physics Publishing, 1999: 96.	1900 kg/m <sup>3</sup> (density)
Jones, Larry. Density notes.	1600 kg/m <sup>3</sup> (density)
The Skeletal System. <i>Oxford Text Book of Medicine.</i> Third edition, Third volume. New York, Medical publications, 1996: 3066.	1000–1200 g/cm <sup>2</sup> (BMD)
Bonnick, Sydney Lou. <i>Osteoporosis, The Hand Book.</i> Third edition. Texas, Cooper Square Press, 2000: 147.	1000 g/cm <sup>2</sup> (BMD)

However this is a very gross approximation since the bones don't have a uniform density. The bone is made of a hard layer and a mineral layer. One of the structures, which makes up the mineral layer of the bone that resembles tubes, is called the Haversian canal. This structure carries organic nourishment required by the bone system. Thin plates called lamellae that contain the marrow of the bone surround the Haversian canal. The yellow marrow is fat and the red marrow is made of tissue that includes the blood cells. The hard layer of the bone is made of collagen. This layer makes up 70% of the density of bone in adults, and 30% in children. The bones next to the inner ear are very rich in these structures and the density is difficult to calculate. Additionally, the simulation assumes zero dampening in the bones. This is another approximation that can affect the results since the fibrous character of bones is guaranteed to cause some dampness to any wave propagation.

The values used for Young Modulus and Poisson ratio are in the table below with some other values for comparison. According to this table using water to fill other subdomains occupied by muscle, fat or brain tissue is a reasonable approximation. This assumption is valid since the muscle and skin densities are very close to that of water

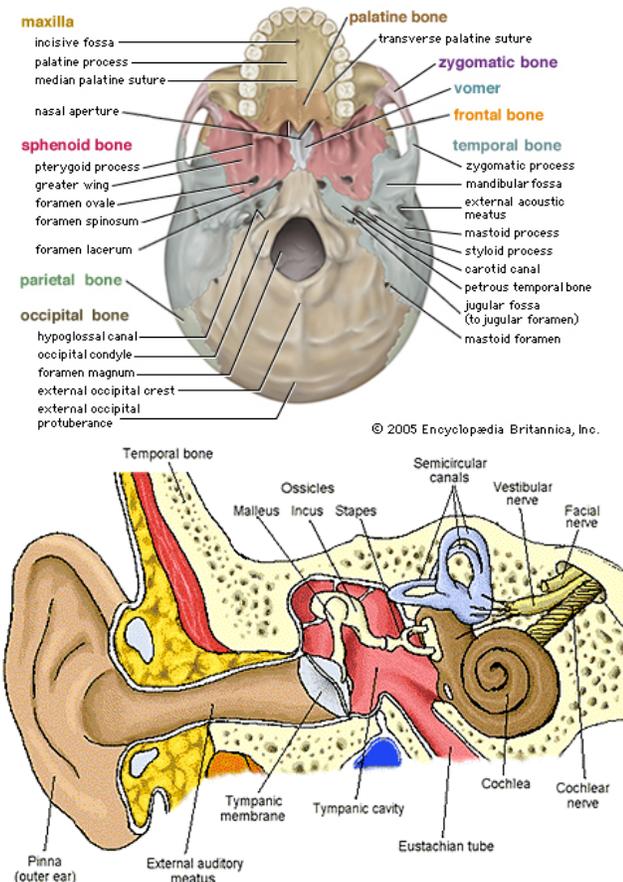
**Table 3. Bone parameters**

Tissue	Graded	Young modulus	Poisson ratio	Density
Fat	Homogeneous	10MPa	0.49914	920
Muscle	Homogeneous	20MPa	0.49876	1070
Bone	2-layered	12GPa	0.30	1260
Titanium	Homogeneous	114GPa	0.326	4480

The cochlea is located deep inside the external acoustic Meatus (the ear canal) inside some bone cavities. The current model is missing all the structure of the inferior part (occipital) of the skull and the structure inside the temporal bone:

- squamosal suture (joint between temporal and parietal bone)
- mastoid process (oval projection behind the ear canal)
- mastoid sinus (air cavity that opens into the middle ear)
- mandibular fossa (oval depression anterior to the ear canal – articulates with mandible)

Additionally, it is missing the mandible, which doesn't seem to have much effect. A model that includes all these complexity is very difficult to implement.



**Figure 4: Internal bone structure and inner ear**

For boundary conditions we use an outer spherical perimeter of the water domain and use the predefined Radiation condition with the Spherical wave option of Comsol. This boundary condition allows a spherical wave to travel out of the system, giving only minimal reflections for the non-spherical components of the wave. The radiation boundary condition is useful when the surroundings are only a continuation of the domain as in the case of underwater environment and we are assuming a large body of water.

To couple the acoustic pressure wave to the solid skull we set the boundary load  $\mathbf{F}$  (force/unit area) on the skull to

$$\mathbf{F} = -\mathbf{n}_s p$$

Where  $\mathbf{n}_s$  is the outward-pointing unit normal vector seen from inside the solid domain.

To couple the frequency response of the solid back to the acoustics problem, use the boundary condition that the normal acceleration equals that of the solid structure.

$$a_n = -\mathbf{n}_a \cdot \left( -\frac{1}{\rho_0} (\nabla p - \mathbf{q}) \right)$$

Here,  $\mathbf{n}_a$  is the outward-pointing unit normal vector seen from inside the acoustics domain.

To find the SIR the sound pressure levels where measured at the source and Cochlea as a function of frequency. The following values where used for the simulation and they directly affect the SIR calculated.

Reference pressure used for the sound pressure level calculation: 1 e-6 Pa

Speed of sound in water: 1500 m/s

Density of water: 997 kg/m<sup>3</sup>

The problem was solved for frequencies from 1 to 20kHz

Finally, the ISTF is just the function that cancels the STF (frequency domain of the SIR). This transfer function was calculated and normalized with MatLab.

## 4. EVALUATION

The model transfer function was evaluated by performing two different experiments on volunteer user subjects. To perform the experiments a prototype of the auditory interface component of the underwater sound localization enhancement device was built. This auditory interface is the part responsible for conveying the sound source location via bone conduction actuators. The configuration selected uses two actuators that are placed in contact with the skin on the cheek bones. The actuators produce sound waves strong enough to impose stresses on the Temporal bones which propagate towards the inner ear and cochlea. The actuators selected for this prototype are "SwiMP3 Bone-Conduction" actuator with a built-in MP3 player and binaural technology. The SwiMP3 actuators have 256 MB of storage and a 10-Hour rechargeable battery which allow us to store a preprocessed sound file and perform the experiments without interruptions. All the preprocessing is performed before hand with MatLab and the model transfer function and a MP3 file is generated.

The first experiment was performed in a normal air environment. The sensation of underwater environment was created by two strategies. Firstly, the sound source signals where attenuated in correlation to the underwater human hearing threshold data. In

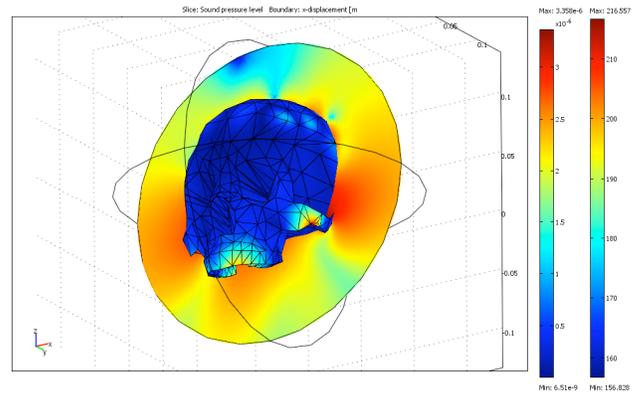
that order the frequencies ranges from 250Hz to 6000Hz are left almost intact (which are the frequencies encountered underwater) and the higher frequencies are attenuated. These signals were processed with the model transfer function and transmitted through bone conduction actuators attached to the subject's skull (cheek bones) using the SwiMP3 device. Secondly, sound cancellation headphones were used to reduce any sound from reaching the drum through the middle ear. Noise-canceling headphones use electronic circuitry to remove noise after it has entered the headphone earcup. Microphones inside each headphone earcup sample the noise field, and an electronic circuit creates an inverse or 'mirror-image' of the noise signal and adds it to the music (or the intended sound). The actual noise and the inverse noise cancel each other out. In this case the intended noise is a typical recording of underwater background sounds at low frequencies. This strategy simulates the underwater sensation of increased wave attenuation for the higher frequencies. Conveniently, sound cancellation headphones are more efficient in canceling higher frequency waves. The higher frequencies are usually attenuated by the headphone body and lower intensities are required to cancel them than for the lower frequencies. The sound canceling headphones selected for the task were the Panasonic RP-HC-55-S which reduces outside noise by 88% (18dB) at 200Hz. This headphones couple with the attenuation of higher frequencies during the preprocessing should ensure that no localization information travels through the middle ear leaving the subject with only the information available through bone conduction.

The second experiment was performed underwater. Its main objective was to verify that the results of the first experiment are correlated with real underwater results. This scheme works since correlation experiments usually require a smaller sample. This experiment is very important because it validates the results from the first experiment. The reason for performing the two experiments is because, given the complexity of the model, the only viable strategy for optimizing it is by performing tests on multiple subjects as the model is further optimized and an underwater test presents some logistic difficulties. The first test, on the other hand, is very easy to perform allowing for larger number of tests and quick turn around and model optimization. For the second experiment the same subject users from the first test submerge in a pool with the SwiMP3 bone conduction actuators in place under the goggles. The pre-recorded (and preprocessed) sound signal is played with the MP3 player while the subject indicates the position with an arm signal. An evaluator records the directions and compares them with the programmed trajectory. After the test the subjects added qualitative comments and a comparison comment with their previous experience

## 5. RESULTS AND DISCUSSION

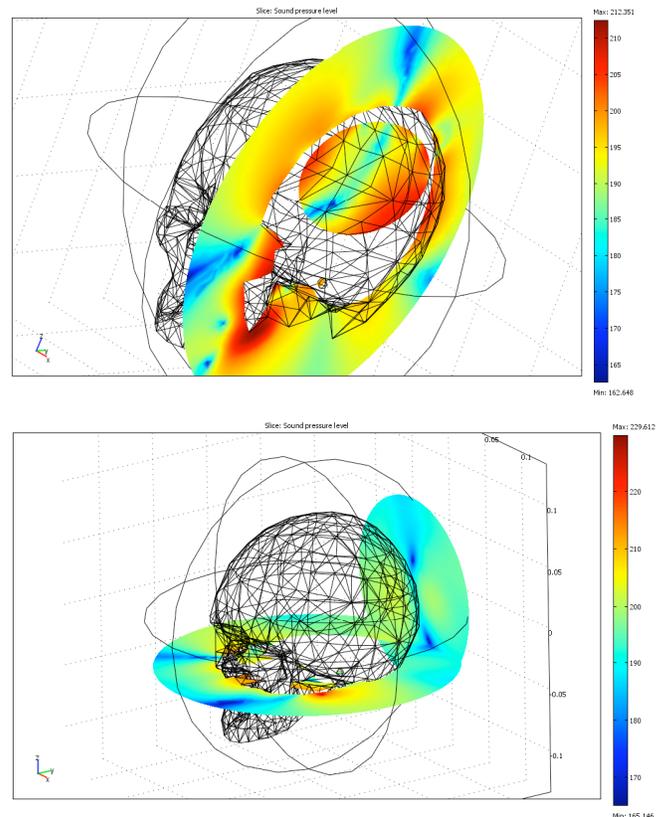
The initial model indicates that the sound propagation of the high frequencies occurs mainly through the temporal bone. The lower frequencies (longer wavelength) modes cannot fit in the small temporal bone channel and propagate through the rest of the skull. This also means that the higher frequencies experience more attenuation and scattering since their channels are more constrained. Unfortunately, these frequencies are the ones that carry most of the position information. The following picture shows two parameters. The slice plot shows the sound pressure level (in dB's) in aqueous subdomain and the boundary plot (skull

surface) shows the deformation of the skull due to stresses caused by the vibrations in the skull. The figure shows the highest deformation (or amplitude of sound wave) at the point of the Zygomatic bone where the source is placed. The source is producing a wave of 1W at 10kHz. The source is at the left side of the skull but there are mirror images of it at the right side and at the back of the head on the right side (not showed). Also notice the high deformation of the left teeth.



**Figure 5: Deformation and sound pressure level plots**

The following plots show the sound pressure levels at the source in front of the Zygomatic bone and at the Cochlea at 10kHz.

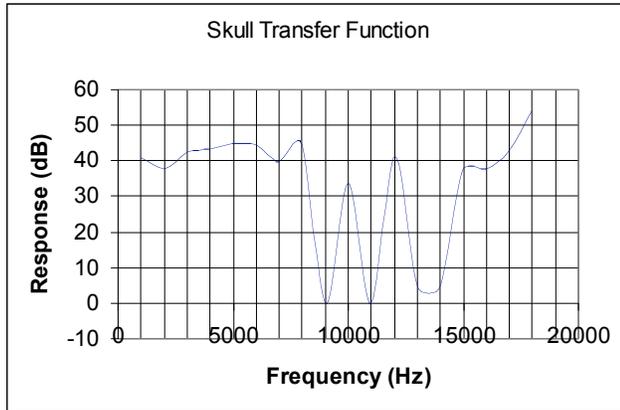


**Figure 6: Pressure sound levels at source and Cochlea**

The equation for the sound pressure level is:

$$SPL = 20 \log_{10} \frac{P_{RMS}}{P_{REF}}$$

The model is very complex and after smoothing the skull CAD model and increasing the maximum mesh dimensions to a third of the wavelength it still takes about 12 hours to compute one data point. For this model 10 data point were calculated at steps of 1kHz. The lack of more data point and the simplified model reduces fidelity accuracy of the STF. The following is a plot of the STF calculated from the sound pressure levels. The ISTF is the inverse of that normalized.



**Figure 5: Skull transfer function**

The model was tested in 6 subjects; 3 females, and 3 males. The This first experiment was performed on the 6 subjects and had the following structure:

1) As many of the following parameters were measured from the subject:

- SEX :
- HEAD\_AND\_TORSO :
- HEAD\_WIDTH\_X1 :
- HEAD\_HEIGHT\_X2 :
- HEAD\_DEPTH\_X3 :
- PINNA\_OFFSET\_DOWN\_X4 :
- PINNA\_OFFSET\_BACK\_X5 :
- NECK\_WIDTH\_X6 :
- NECK\_HEIGHT\_X7 :
- NECK\_DEPTH\_X8 :
- TORSO\_TOP\_WIDTH\_X9 :
- TORSO\_TOP\_HEIGHT\_X10 :
- TORSO\_TOP\_DEPTH\_X11 :
- SHOULDER\_WIDTH\_X12 :
- HEAD\_CIRCUMFERENCE\_X16 :
- SHOULDER\_CIRCUMFERENCE\_X17 :

With these parameters a suitable HRIR out of the 51 available was selected. Each data set has a card with all this data from the original subject measured.

2) The subject tested the HRIR with the noise canceling headphones. The maximum resolution was 15 deg azimuth and 15 deg elevation. The sound swap followed 2 trajectories:

- Horizontal in a 360 deg range
- Vertically in a 45 deg range

The user had the opportunity to indicate the direction of the sound and comment in perceived resolution. The subject also specifies if externalization had been achieved. These results are the benchmark for the following tests.

4) The same test was performed with the HRTF modified by the ISTF. The subject performed the same test as before and additionally commented about the quality of the perceived spatialization compared with the first experiment.

For this experiment only one out of six subjects reported elevation perception with a resolution of 90 deg or a sound in plane or at the zenith. None of the subjects reported externalization. The resolution reported for the azimuth was 20 deg. However, the direction indicated during the experiment indicated a 15 deg resolution in agreement to the trajectory programmed. The reported results for both sets of experiments (only HRTF and HRTF \*ISTF) are identical but the subjects reported a more clear perception with second set (HRTF \*ISTF).

The second experiment was performed underwater with only four subjects; 3 males and one female. Its main objective was to verify that the results of the first experiment are correlated with the underwater results. The experiment was structured as follows:

1) As many of the following parameters were measured from the subject The subject submerged with the bone conduction actuators and a pre-recorded sound swap with a different trajectory (only horizontal and vertical) and indicated the position with and arm signal. An evaluator recorded the directions. After the test the subject was allowed to add qualitative comments as externalization, etc.

2) The experiment was performed with the two sets (HRTF only and HRTF\*ISTF) to compare with the last experiment.

For the first set all the subjects reported results identical to the first experiment with the first set. For the second set (HRTF\*ISTF) all the subjects reported a higher fidelity and more clear perception of sound localization in the underwater experiment than in the air environment experiment. The resolution reported was 15 deg in agreement with the direction indicated with their arms. One of them reported elevation, the same one that reported elevation previously, with a resolution of 90 deg. None of them reported, externalization.

One possible reason for the agreement of the resolution reported and the indicated during the experiment is that in the underwater experiment the subject is more focus on the sound and has less distractions. This could also be an indication why the fidelity is higher in the underwater experiment. Additionally, in the underwater experiment some sound waves travel through the middle ear improving the perception of the higher frequencies and the clarity of the sound. From the results we can conclude that the air experiment agrees with the underwater experiment and is a

good evaluation tool. This conclusion is however, as good as the data collected. Since only 6 subjects were tested there a good reason to doubt the validity of these results. The second conclusion is that we have insufficient data to declare whether or not the HRTF\*ISTF is better than only the HRTF. This is due to two reasons: the quality of the HRTF used as reference didn't prove to be very good and it was not possible to reach a good perception of externalization and elevation perception with it. The second reason is that the quantitative results don't indicate any improvement despite the fact that all the qualitative responses indicate improvement.

## 6. CONCLUSION

The initial model didn't show any qualitative improvement over the simple HTRF with respect to underwater sound spatialization. Quantitatively, every subject reported improvement in the quality of the sound perceived. One reason for the ambiguity of these results is the suspected bad quality of the HRTF used as reference. All the subjects reported a good agreement between the test performed on an air environment and the underwater tests proving that this could be a good evaluation tool. However, given the small set these results should be taken with a grain of salt. The model is very simplified and any improvement will require a increase in resource available and specifically computational power.

## 7. REFERENCES

- [1] Davies, T. C., 2008. Ultrasound for Human Echolocation. PhD Thesis University of Waterloo.
- [2] Norman A, N., Phelps, R., Wightman, F., 1970 Some Observations on Underwater Hearing. *Journal of Acoustic Society of America*, vol50, n2.
- [3] Parvin, S. Limits for Underwater Noise Exposure of Humans Divers and Swimmers. Subacoustech. Underwater research and consulting.
- [4] Yu Eriguchi., et all 2008 System Design and sound localization System of an Autonomous Underwater Robot "Aquabox" 11th International Autonomous Underwater Vehicle competition Journal, pp. Kit-1 to KIT-9
- [5] <http://recherche.ircam.fr/equipements/salles/listen/index.html>
- [6] Gaver, W., Auditory Interfaces, Royal College of Art Kensington Gore
- [7] Hollien, H., 1973 Underwater Sound Localization in Humans *Journal of Acoustic Society of America*, vol55, n5.
- [8] Pec, M. 2007 Personalized Head Related Transfer Function Measurement and Verification through Sound Localization Resolution, Eurasip
- [9] Picinali, L. Techniques for the Extraction of the Impulse Response of a Linear and Time-Invariant System