Enabling Humans to Hear the Direction of Sounds Underwater - Experiments and Preliminary Results

Teong Beng Koay, Jolyn Tan, Soo Pieng Tan, Soon Keat Yeo, Harold Tay, Mandar Chitre, and John Robert Potter

Acoustic Research Laboratory, Tropical Marine Science Institute
National University of Singapore, 12A Kent Ridge Road, Singapore 119223

koay@arl.nus.edu.sg

***Abstract* - Sound localization by the human auditory system is known to be ineffective underwater. In a research study at the Acoustic Research Laboratory, we introduce additional directional cues from ultra-sonic frequency components. This enables us to work with smaller wavelengths and build simple, small, directional receivers in order to introduce Inter-aural Temporal Differences (ITD) and Inter-aural Intensity Differences (IID) cues to the human subjects. Along with a real-time acoustic bandwidth compression algorithm, we are able to introduce a sense of direction into their hearing in water.**

The prototype system consists of a pair of directional receivers of less than 60 mm diameter that are spaced and angled in such a way that high frequency directional cues from 20 kHz-200 kHz are created in the audio hearing range, providing a sense of direction.

Several experiments have been conducted in a small tank to study the performance of the system. The preliminary results are very promising. Independent tests on the directional receivers show that subjects are able to differentiate direction to within ±12 degree 95% of the time. The subjects claim that there was a clear sense of direction in their hearing tests.

Results from swimming pool experiments indicate that subjects wearing the prototype system demonstrate capability in direction localization, although the sense of direction seems somewhat reduced compared to the independent tests on the directional receivers. This paper presents the experimental design, and preliminary analysis of the data.

I. INTRODUCTION

The ability of the human auditory system to sense direction in water is known to be extremely ineffective. The inefficiency of the human auditory system when operating in water and the higher sound speed in water are responsible for significant reduction in ITD and IID cues for auditory perception. Additionally, the relatively large wavelengths and differences in the auditory pathways (e.g. bone conductivity in water as compared to outer ear canal in air) have rendered the human direction perception mechanisms ineffective underwater. Although some experiments [1] have showed that human subjects were able to localize sound underwater to some extent, sound localization in water is poor if not impossible for most people. This not only makes the acoustic perception of underwater sounds monophonic and less exciting compared to stereophonic sounds, but also greatly reduces safety margins in the recreational diving community as they are not able to localize dangers that are associated with sound (such as high speed boats passing by).

Many underwater ambient noise sources contain ultrasonic frequency components (above 20 kHz). Examples of such noise sources include reefs with biological noises produced by snapping shrimp [2,3]. Snapping shrimp noise contains significant spectral energy well above 100 kHz. Boat noise from small, high-speed marine craft may contain high-frequency acoustics components with acoustic source levels around 140 dB re 1μPa @ 1m at 30 kHz [4]. Dolphins and some species of whales [5] also produce ultrasonic clicks.

Acoustic reflectors and lenses with a relatively small aperture can be used to focus sound onto a single receiver to discriminate the direction of these high frequency acoustic components. If the bandwidth can somehow also be down-shifted into the human audio range and presented to the user in a representative way, this could enable human subjects to perceive an enhanced sense of direction in addition to the audio band acoustics they already perceived.

Being able to introduce a sense of direction into human hearing in water would open up a new domain of sensation and enhance the safety of the diving community. It would be possible to create a more exciting underwater exploration experience by allowing the user to correlate sound to visual observations such as coral reefs, fish or boats. This is a similar facility to which we enjoy in air.

To create the receiving aperture, a pair of directional receivers are angled and spaced in such a way that they introduce IID and ITD that is comparable with that created by human auditory perception in air. The ability to down-shift the ultra-sonic frequencies into audible range is achieved using an Acoustic Bandwidth Compression (ABC) algorithm [6]. The system is configured to compress signals between 20 kHz and 200 kHz into audible frequency band from 1.7 kHz to 17.7 kHz, providing a bandwidth compression of about 11.34:1. The output sounds are played through a pair of stereo underwater headphones to the human subject.

The directional receivers were first evaluated for their capability to provide a sense of direction with the receivers placed in a tank and the human subjects wearing a pair of headphones in air. This was then followed by underwater tests where four divers, one at a time, were placed in a seat in the centre of an aluminium structure that held five acoustic transmitters arranged on a semi-circle of 1m radius. An acoustic signal was randomly played from one of the five transmitters and the subjects were required to identify the direction of source. This was repeated multiple times. The answers were recorded and checked against the actual position of the test transmitter.

Preliminary experimental results suggested an angular bias for direction perception. This bias was believed to be a combination effect from the system as well as auditory system of individual subject. A calibration process at the beginning of the underwater experiment for each subject helped correct this bias in subsequent experiments.

II. CONCEPTUAL DESIGN OF SYSTEM

The prototype system consists of a pair of directional receivers, a dual channel preamplifier, a USB National Instrument multifunction analogue input and output module, a computer, and a pair of underwater headphones.

Each directional receiver provides a -3dB beamwidth of 45 degrees at 60 kHz. They are placed close together, angled away from each other, somewhat akin to the placement of human ears. The spacing between the receivers is about 100 millimetres, chosen so that they produce ITD (after signal compression using the ABC algorithm) similar to human perception in air.

The computer is used to digitize the incoming acoustic signals, run the ABC algorithm and produce audible output in real-time to the underwater headphones. The input signal is sampled at 500 kSa/s and the output audio is sampled at 44.1 kSa/s.

The directional receivers are mounted near the cheek of the subject, while the headphones are placed over the ears, allowing the system to move with the head.

III. TANK TEST OF THE DIRECTIONAL RECEIVER

A number of mechanisms could affect the sensing of direction if the listener were in water. As an example, the high sound speed in the skull [7,8] could transfer acoustic signal played in one ear to the other almost immediately. Although the coupled signal would be modified by the bone's transfer function and produce IID, creating a second arrival might confound the ITD. With these potential confusions, it is important to study the design of the receiver alone, in isolation from other possible causes that could affect the sense of direction, in order to arrive at a proper receiver design.

We therefore evaluated direction finding performance using the directional receivers alone in a test tank with the human user seated in air.

A. Tank experiment setup

The experiment was conducted in a small tank measuring 2.5 x 2 x 2 meters. The source and the receivers were mounted at the mid-water depth (1m) in the tank in order to separate surface/bottom reflections from the direct arrival as much as possible. Figure 1 shows the experimental setup.

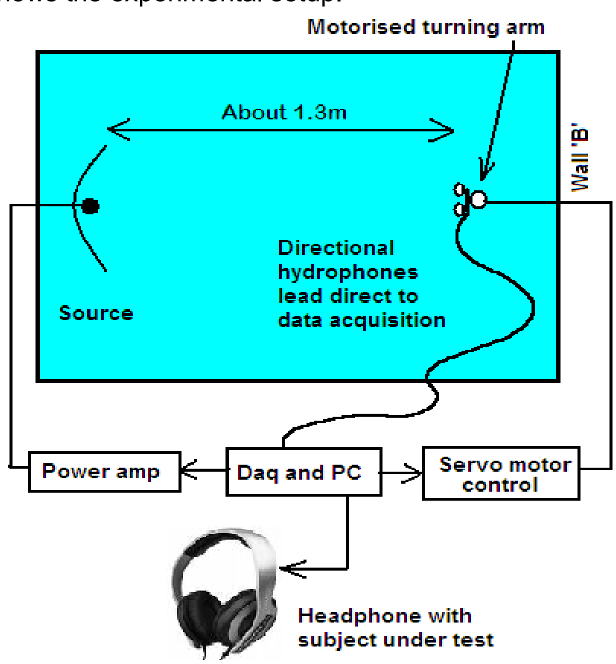

Figure 1: Schematics of the tank experiment

The source was mounted on a parabolic reflector dish in order to minimize reflections from the tank walls by directing most of the acoustic power directly towards to receivers while minimizing the acoustic power radiated to the side walls, surface or bottom. The main reflections anticipated are those from wall 'B' in Figure 1, which is directly in front of the source. This might potentially cause a 180 degree front-back ambiguity for the receiver during the experiment.

The source was placed as close as possible to the wall in order to maximize the source-receiver distance. The aperture of the transmitting parabolic dish was about 450 millimetres, chosen to be larger than the spacing between the receivers (100 millimetres). This was to ensure that the receivers are always in the main beam of the transmitter in all of their turning angle positions.

The receiver pair was placed about 1.3 meters away from the source, which was the furthest possible allowed by the tank setup. The receiver pair was mounted on a pole equipped with servo-motor that can be remotely rotated to an accuracy of 0.5 degree. During the experiment, test software randomly rotated the heading of the receiver pair to an initial orientation at the beginning of the experiment while the subject's task was to manually rotate it until he/she thinks the sound was coming from directly in front. The tank was not visible to the subject.

The test signals are generated by a PC and amplified by a power amplifier before being fed into the source transmitter. Along with a data acquisition module, the PC digitizes the received acoustics from the receiver pair, processes them using the ABC algorithm and plays the resultant audio to the subject under test through a pair of high quality headphones. The entire process is performed in pseudo-real-time with a small latency.

A user interface for the subject allows remote control of the heading of the receivers with simple mouse clicks. The servo control system is set for low rotation speed so that the subject does not over-shoot the heading control relative to his/her hearing perception, as the processed feedback contains a small delay.

B. Selection of test signal

The test signal frequency band was selected based on frequency components that are common to acoustics from natural biological (snapping shrimp dominated reef noise) and boat noise in a typical coastal ambient noise environment. Frequencies used by marine mammals (echolocation clicks from dolphin and whales) were also of interest.

Since we are interested in investigating the sense of direction in underwater human hearing, the test signal was designed to be long enough for the human auditory system to perceive while small enough to operate in a small tank with minimal multi-path overlap. The integration time of the human auditory system is approximately 25 milliseconds [9], which corresponds to a length of 37.5 m assuming a sound speed of 1.5 km/s. This implies that in almost all practical situations of the tank experiment, there is to be some multipath overlap to contend with.

Since a parabolic reflector was used at the source to limit the side reflections, the frequency of the test signal was selected so that the wavelength is to be at least 10 times smaller than the aperture size of the transmitter. With the above considerations, we loosely selected the test signal to be 60 kHz CW pulse with pulse length of 40 milliseconds. The 25-millimetre wavelength of 60 kHz signal allowed us to have an aperture-to-wavelength ratio of 18:1 with the source reflector so that nearly plane-waves would be generated. We used a pulse length of about 450 milliseconds (after bandwidth compression), ensuring the signals could be clearly heard by the subjects. These test parameters were used in all experiments, both in the tank and the swimming pool, so that the results could be compared.

C. Results from the tank experiment

Four subjects carried out the tank test with test repetitions varying from about 20 to 30 (spread over multiple sessions), with the source located at a bearing of 90 degrees referenced to the angular origin of the experimental setup. A total of 91 tests were carried out and the results for each subject are shown in Figure 2. Most of the subjects were able to estimate the direction with a standard deviation of less than 5.6 degrees. Subject 4 initially had a larger standard deviation of about 10 degrees. This is considered an anomalous result because one of their estimates was 35 degrees away from the mean estimates. This data point was subsequently ignored in the standard deviation computation and the standard deviation estimate for this subject then reduced to 5.6 degrees.

It is also seen that the mean of the heading estimates from individual subjects vary between 75 degrees to 85 degrees. This indicates a bias of 5-15 degrees from the true source location of 90 degrees. Further examination shows that all the biases are deviated to the left, strongly suggests that this was a systematic error in the experiment setup. Our hypothesis for this observation was that the variations of overall channel gain between left and right signal paths caused an incorrect IID cue to be presented to the subject. This could be due to variations of the sensitivity of the hydrophones, component tolerance in the signal conditioning electronics circuits, or variations in the gain of the playback mechanism. A second set of experiments were conducted to test this hypothesis.

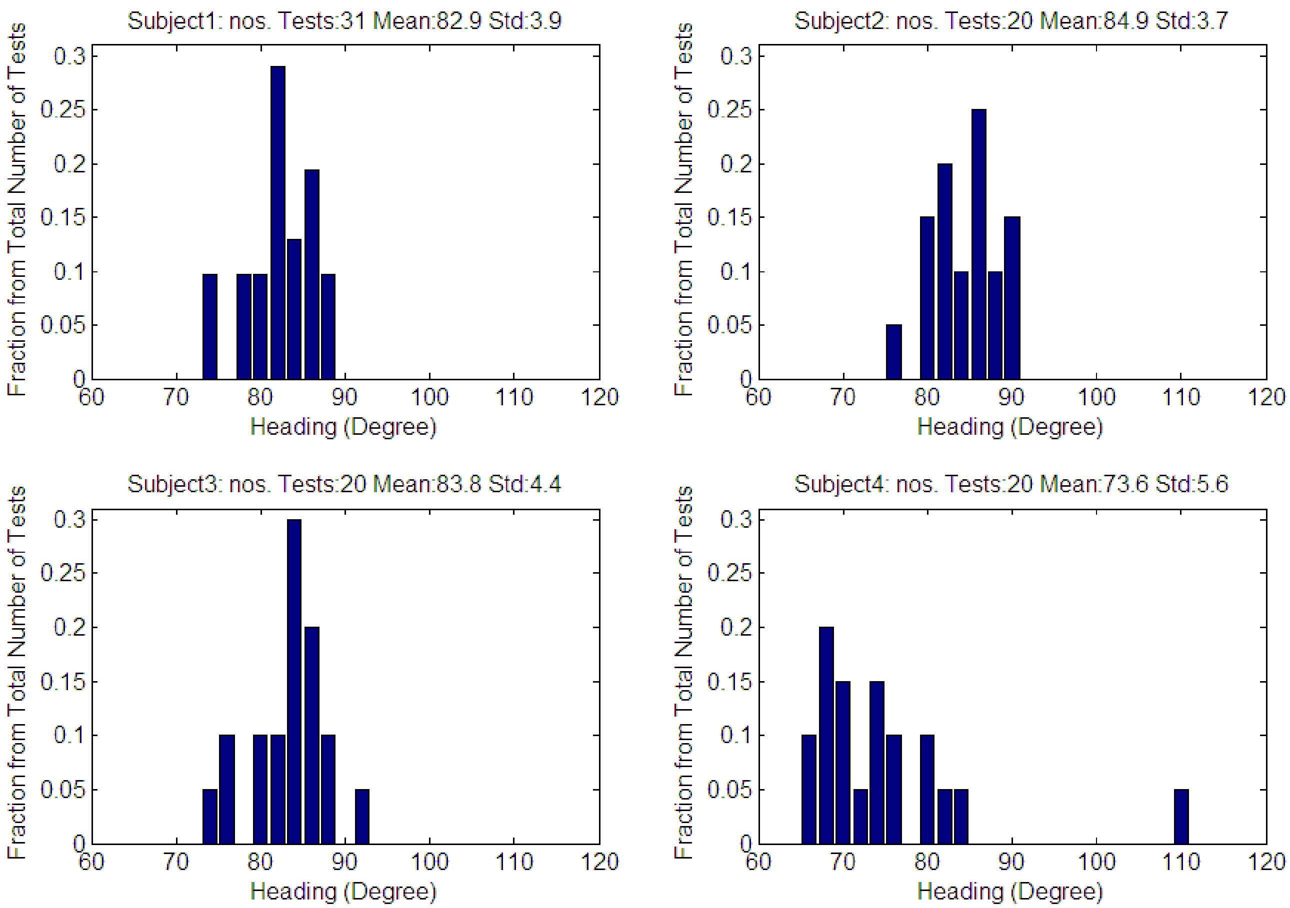

Figure 2: Distribution of headings estimates by individual subjects of a source placed at 90 degree. Direction perception of subjects are not calibrated

For the second tank experiment, the setup was first calibrated to compensate for the left-right channel gain deviation at the beginning of the experiment. Since the sense of audio direction is given by both ITD and IID, changing the ratio of the left and right electronics gain shifts the perceived direction. This was used to compensate for the observed deviation. The receiver pair was first rotated to face the source and the subject digitally adjusted the gains of left and right channels until subject felt that the source was directly in front. This equalized the overall gain of the entire experiment setup, including the left-right sensitivity deviation of the auditory system of the subject.

The experiment was then repeated with three subjects with a total of 91 tests. The results of the experiment are presented in Figure 3. The angular biases observed in previous experiment were eliminated. The mean of heading estimates are very close to the direction of the source (within 2 degrees), and the mean heading estimates among the subjects are within 3 degrees from each other. We therefore concluded that the hypothesis was correct and similar steps can be taken to equalize the gain variation of left-right channels in future experiments.

Another interesting observation is that Subject 2 had registered bearing 266 degrees as one of the answers although the source was located at bearing 90 degrees, see arrow notation in Figure 3. This is consistent with the anticipation that there could be front-back ambiguity caused by the experiment setup as the source signal was reflected by wall 'B' denoted in Figure 1. The reflected signal from wall 'B' will appear as a phantom source located at a bearing 270 degrees although with lower acoustic intensity. Another possible reason could simply be the front-back ambiguity of the direction perception based on ITD component. The data point at 266 degree was ignored in the subsequent analysis.

This was a rare mistake perhaps because the randomly generated initial positions of the receiver-pair were always limited between 0-180 degrees. Hence the heading estimation can normally be expected to converge toward the actual source that is located at 90 degrees rather than its reflected image.

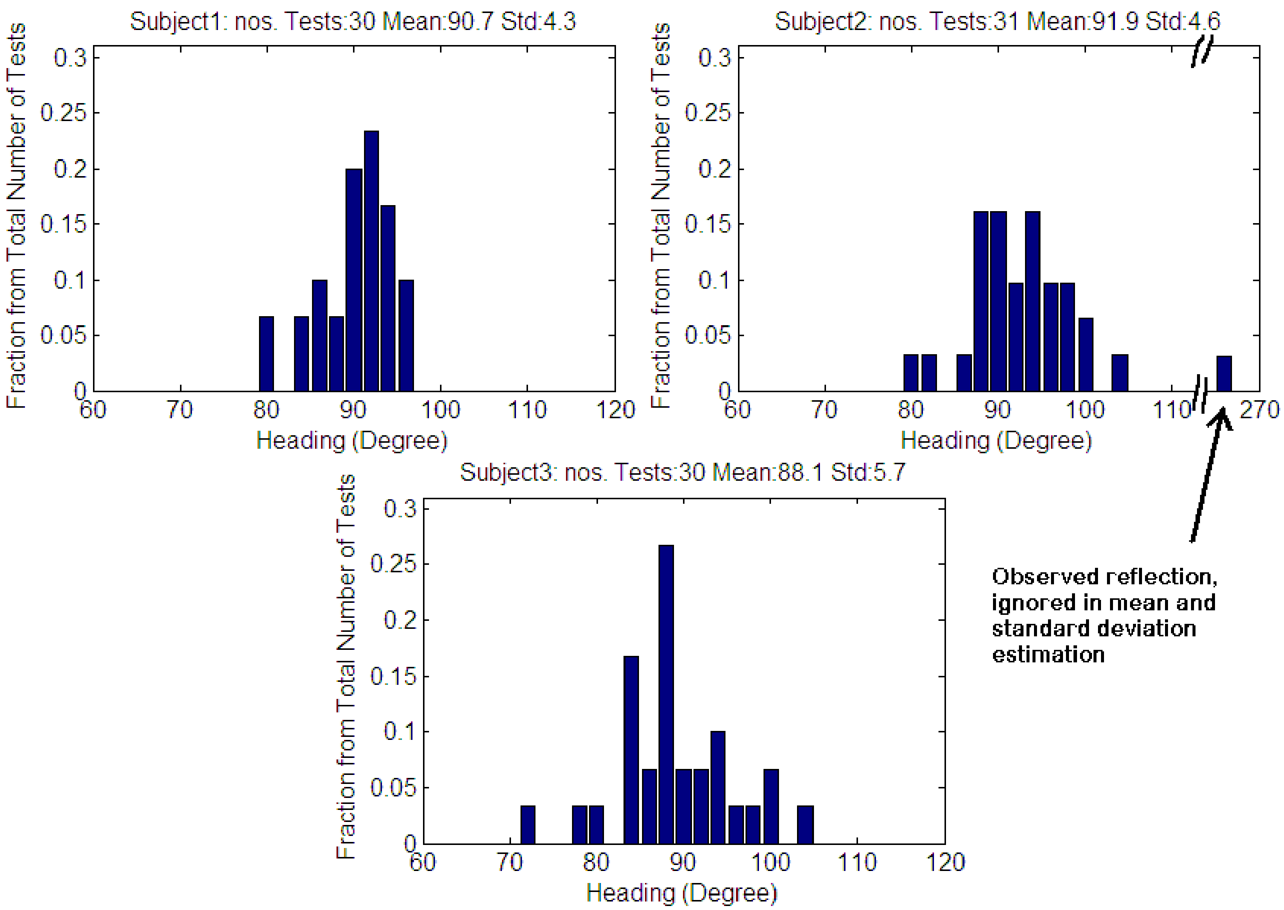

Figure 3: Distribution of headings estimates by individual subjects after the direction perception of subjects is calibrated. Source is placed at 90 degree

Figure 4 shows all the estimates of source location from all subjects. The mean estimate is 91 degrees with a standard deviation of less than 6 degrees. A Lilliefors test at 5% significance level shows that the heading estimates are normally distributed at the 95% confidence interval, with p-value of about 0.2. We conclude that the heading estimation is accurate to better than 6 degrees (one standard deviation) about 67% of the time and less than 12 degrees (two standard deviations) 95% of the time.

IV. SWIMMING POOL TESTS OF THE DIRECTIONAL RECEIVER

The experimental setup in the swimming pool is illustrated in Figure 5. The subject was securely positioned in the middle of the setup throughout the experiment.

The right portion of Figure 5 shows the underwater fixture designed for this purpose. The fixture is about 2 m high yet it can be easily deployed from the side of the pool with the aid of swimmers. The fixture contains five transmitter holders with adjustable spacing located in a semi-circle placed 1 m away from the subject.

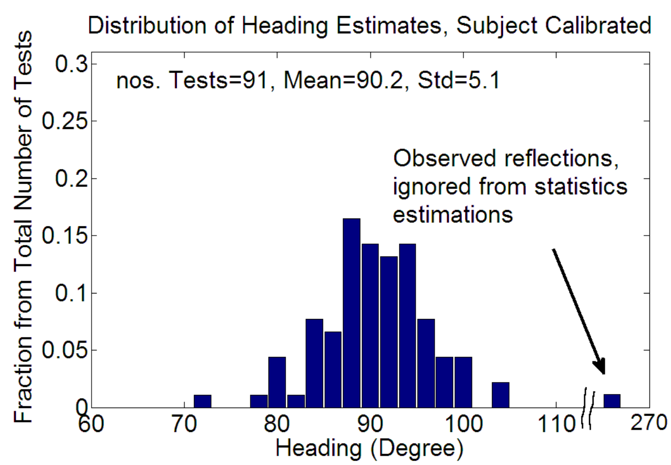

Figure 4: Histogram of estimated angular locations of source located at bearing 90 degree from all subjects

A sitting platform and a footrest allow the subject to position him/herself at equal distance from all the transmitters. An underwater CCD camera is mounted on top of the fixture to film the experiment. A flexible mount is provided to hold the electronics response panel so that the hands of the subject are free throughout the experiment.

The electronics setup uses the same prototype directional receiving unit as was used in the tank experiments as described in section II.

In addition to the prototype unit, the setup contained a separate electronics module to automate and coordinate the entire experiment. The module transmitted a selected test signal from one of the five transmitters selected by random. It also provided visual indications through the LEDs located on the electronics response panel to inform the subject about the various stages of the experiment. The subjects would provide their estimates of the transmitter they thought was the source using the same panel. The automation module recorded the responses from the subject as well as the test conditions and parameters of each test. They were then cross-checked to tabulate the performance of the subject. A safety swimmer was always in the water in case of need. The left portion of Figure 5 shows the overall schematic of the setup.

V. RESULTS FROM SWIMMING POOL TESTS

Two sets of swimming pool experiments were conducted with 4 subjects. The first experiment was conducted with the sources located at an angular separation of 45 degrees, while the second experiment reduced the angular separation to 25 degrees to investigate resolution performance. Both experiments employed the same signal parameters used in the tank test (see section III) for the generation of test signals. Subjects were free to move their head and body as long as they did not leave the sitting support, i.e. their position relative to the transmitters were preserved.

A. 45 degree spacing tests

A total of 200 tests were carried out in first experiment spread across 4 subjects. The positions of the sources labelled 1 to 5 corresponds to -90 degrees, -45 degrees, 0 degree, 45 degrees, and 90 degrees respectively. The 0 degree source was located directly in front of subject in their normal sitting position. The subjects gave 170 responses within the allocated time (35 seconds) for each test. The remaining 30 tests were declared void since no response was given within the time limit.

TABLE 1 tabulates all 170 transmitter locations against the responses given by the subjects in the form of percentage of total count of valid responses. The bold diagonal column would contain the only non-zero elements if all of the responses were correct. In this case the spread of the responses into other columns gives a sense of how the errors were distributed. The result is defined as correct only if the estimated source location matches the actual source location for that particular test.

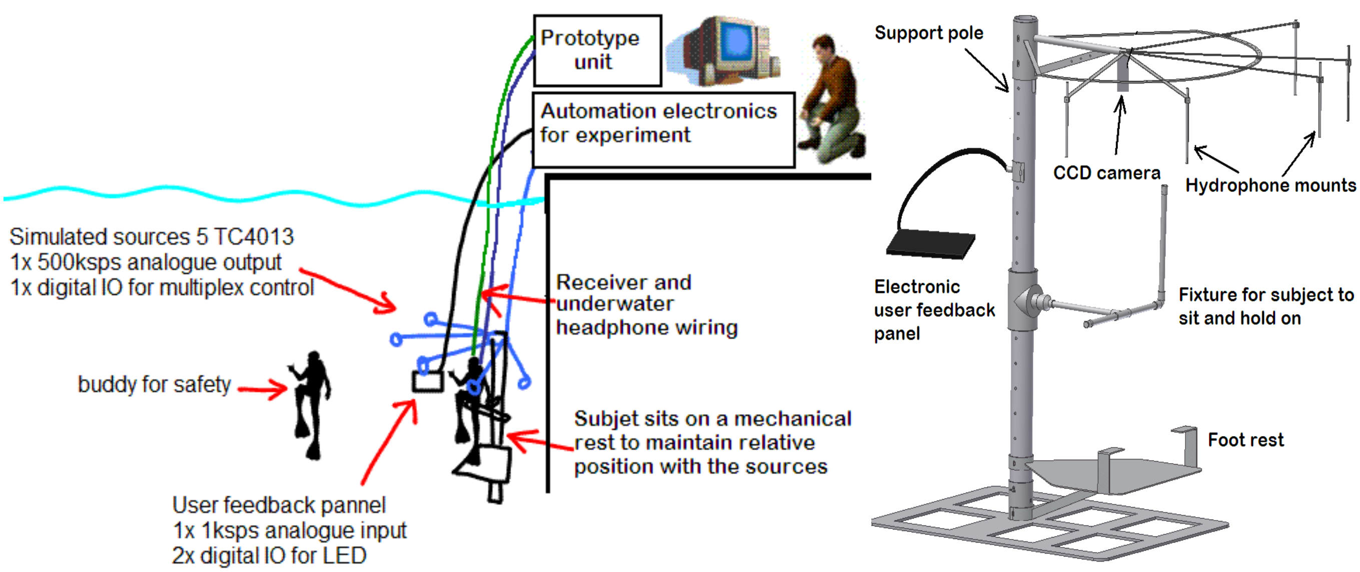

Figure 5: Experiment setup of swimming pool experiment

TABLE 1: PERCENTAGE DISTRIBUTION OF ESTIMATES FROM SUBJECTS VERSUS SOURCES LOCATION (45 DEGREE SPACING)

		Source position				
		1	2	3	4	5
Subject's response	1	**18.8**	0.6	0.6	0.0	0.0
	2	0.0	**14.7**	4.1	0.6	0.6
	3	0.0	4.1	**15.9**	0.0	0.0
	4	0.0	0.0	1.2	**15.9**	2.9
	5	0.0	0.0	0.0	1.2	**18.8**

The results shown in TABLE 1 show that the subjects were able to successfully identify the source locations about 85% of the time. The largest number of errors were made discriminating between source positions 2-3 (between -45 degree and 0 degree), with 8.2% of total responses. The next most common error (4.1%) was confusion between source positions 4-5 (45 degree and 90 degree to the right), followed by source positions 3-4 (1.2%) and positions 1-2, 1-3, 2-4 and 2-5, with each 0.6% of responses (or 1 mistake).

This suggests that signals coming directly from the front pose greatest challenge to be detected in this experiment. This could be partly explained by there being more choices when the source was transmitted from directly in front. However, the majority of the errors were to be one step away from the correct answer.

B. 25 degree spacing tests

The second experiment was carried out with the sources positioned at 25 degrees angular separation. The five sources were again labelled 1 to 5; located at -50 degrees, -25 degrees, 0 degree, 25 degrees, and 50 degrees respectively with 0 degree being directly ahead of the subject. A total of 250 tests were conducted across 3 subjects. 100% of the estimates were considered valid with each of the tests answered within the designed 35 seconds time limit. TABLE 2 summarizes the results from the second experiment set in the form of percentage of total number valid tests.

Subjects were able to correctly detect the source locations 76% of the time. Most errors were made in discriminating between sources located at position 4 and 5 (between 25 degree and 50 degree for this test), about 8.8% of the total responses. The rest of the mistakes were confusions between the source and its adjacent choices, with the number of mistakes evenly distributed across all directions, forming 4.8%-5.2% of total tests at each angle.

TABLE 2: PERCENTAGE DISTRIBUTION OF ESTIMATES FROM SUBJECTS VERSUS SOURCES LOCATION (25 DEGREE SPACING)

		Source position				
		1	2	3	4	5
Subject's response	1	**16.8**	3.2	0.0	0.0	0.0
	2	1.6	**16.4**	2.0	0.0	0.0
	3	0.0	3.2	**13.6**	3.2	0.0
	4	0.0	0.0	1.6	**15.2**	3.2
	5	0.0	0.0	0.0	5.6	**14.4**

C. Performance variations between 45 degrees and 25 degrees source spacing experiments

Figure 6 shows the probability of successful direction estimation across the different angles from both sets of experiments. Each bar is calculated as a percentile of successful localization at a particular direction from the total tests issued in that direction ignoring the tests without valid responses. The x-axis shows the position labels of each source, indicating the choices available to the subjects. Note that the angular spacing between the adjacent positions is 45 degrees and 25 degrees respectively.

It is observed that the probability of successful direction detection towards the edges (positions 1 and 5) is slightly higher (about 15%) than the rest, possibly because there was no available choice at greater angle in each of these cases. The probability of successful direction detection from the 25 degrees source spacing experiment shows a slight tendency to decrease from source location 1 to 5, but the differences are statistically insignificant.

In general, the performance of direction localization with a 25-degrees source spacing is slightly lower than for the 45-degrees source spacing experiment. This matches the expectation that the localization of acoustic direction will get more difficult when the source angular spacing gets smaller. Nevertheless, the differences are small. Along with the high success rate in estimating the source locations (about 68-84%) of all locations at 25-degree source spacing, the result suggests that we are still far away from the limit of angular resolution capability of the setup.

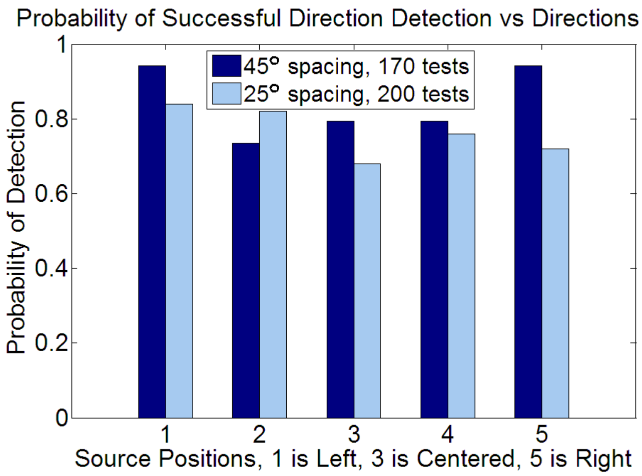

Figure 6: Comparison of probability of detection at different source locations

Figure 7 summarizes our observations as the probability of occurrence for each possible distance of error. The distance of error is defined as how far an estimate is from the actual source location, in steps of source spacing. The x-axis shows distance of error. A correct response has zero distance of error. Responses that are one choice to the right of the actual source location are labelled as 1 and so on. The bars show the probability of occurrences (number of user responses/number of possible occurrences) for that particular error distance.

Figure 7 shows that there is a very low probability that subjects will make errors that are more than one spacing step away from actual source location. Over 90% of the total errors from the 45-degrees-spacing experiments and 100% from 25-degree-spacing experiments are within one spacing step from the actual source location.

An interesting observation is that the errors made in the 25 degrees source spacing experiment were less spread out compared to errors from 45 degrees source spacing experiment. This is contradictory to the intuitive expectation that a higher number of large errors would be made as the angular spacing becomes closer. This could be due to the fact that the subjects had gained more experience over time, as they had performed a number of 45 degrees source spacing tests before the 25 degrees spacing experiment. This suggests that with experience, subject could perform better over time.

If this is the case, then the observed minor performance drop between the 45-degrees and 25-degrees source spacing experiments could actually be a sum of a larger true resolution-related performance drop and a learning performance improvement. This is because the performance gain due to increased experience was included in the observation. Further testing at the 45 degree spacing would be needed to test this hypothesis.

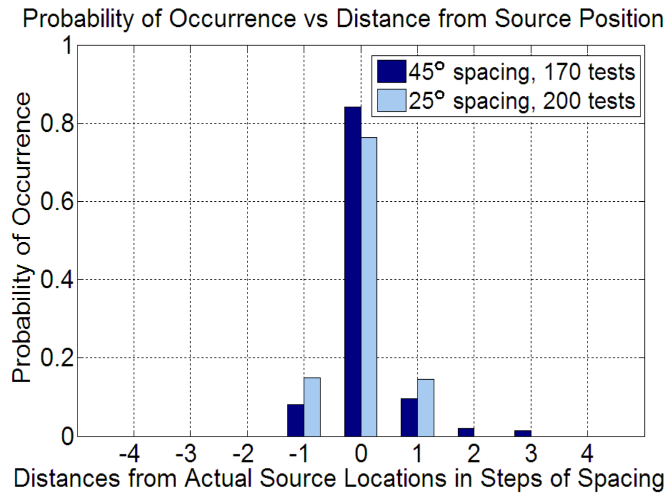

Figure 7: Probability of occurrences and distance of estimates from source position

VI. CONCLUSIONS

We have presented a simple methodology to improve the sense of direction of the human auditory system underwater for acoustics containing high frequency component. This is done by extracting time and amplitude cues from high frequency components of the acoustic soundscape in pseudo-real-time. We have built a prototype system to test this idea. The system consists of a pair of directional receivers formed by parabolic reflectors focussing sound onto single receivers, followed by a bandwidth-compression algorithm that provides IID and ITD cues via headphones to the human auditory system.

The preliminary results show that our prototype system provides the user with cues that permit better than 12 degrees resolution 95% of the time. A human subject experiment in a swimming pool showed that subjects were able to identify the locations of sources spaced at 45 degrees apart 84% of the time and 76% of the time when sources were spaced 25 degrees apart.

It appears that we have yet to test the subjects to the limit of the angular resolution for this methodology, and/or subjects are able to significantly improve their performance with training.

It therefore seems possible to provide marine scientists, biologists and recreational dive community with an effective sense of direction in underwater hearing using only a simple portable device. When combined with the visual sense, this additional sensory dimension would allow them to appreciate and explore underwater world much more effectively and with far greater appreciation with stereo sounds in water.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to NUS Sports and Recreational Centre for always being supportive in making the swimming pool available for our research.

The authors would also like to thank Dr. Lum Wai Hoong for providing a set of underwater hydrophones and developing a pair of underwater audio-playback mechanism for the experiment.

REFERENCES

[1] H. Hollien, “Underwater sound localization in humans,” J. Acoust. Soc. Am. 53, 1288, 1973

[2] W. W. L. Au & K. Banks, “The acoustics of the snapping shrimp Synalpheus parneomeris in Kaneohe Bay,” J. Acoust. Soc. Am. 103: 41–47, 1998.

[3] T.B. Koay, E.T. Tan, M. A. Chitre & J. R. Potter, “Estimating the spatial and temporal distribution of snapping shrimp using a portable, broadband 3-dimensional acoustic array,” Oceans 2003 Marine Technology and Ocean Science Conference (MTS/IEEE), San Diego, USA, September 22-26, 2003

[4] B. Kipple and C. Gabriel, “Glacier Bay watercraft noise – Noise characterization for tour, charter, private, and government vessels,” Technical report for Glacier Bay National Park and Reserve, 2004.

[5] W.L. Au, “The Sonar of Dolphins,” Springer-Verlag, 1993

[6] T.B. Koay et al., “A compact real-time acoustic bandwidth compression system for real-time monitoring of ultrasound,” in Proc. Oceans 2004, Kobe, Japan, November 2004

[7] Kang Il Lee, Heui-Seol Roh, and Suk Wang Yoon, “Acoustic wave propagation in bovine cancellous bone: Application of the Modified Biot–Attenborough model”, J. Acoust. Soc. Am. 114, 2284, 2003

[8] F. J. Fry, and J. E. Barger, “Acoustical properties of the human skull”, J. Acoust. Soc. Am. 63, 1576, 1978

[9] A. Azzali, A. Bellini, A. Farina, and E. Ugolotti, “Design and implementation of Psychoacoustics Equalizer for Infotainment”, DSP Implementation Day, Politecnico di Milano, 23 September 2002