Measuring the augmented sound localization ability of humans in the underwater environment

Koay T.B., Yeo S.K., Tan G.H., Tan S.P., Seekings P.J., Chitre M.
ARL, Tropical Marine Science Institute
National University of Singapore
18 Kent Ridge Road, Singapore 119227
Email: koay@arl.nus.edu.sg

***Abstract*—Humans are poor, if not incapable, at localizing sound underwater due to significant reduction in Inter-aural Temporal Differences (ITD) and Inter-aural Intensity Differences (IID) caused by reduced impedance mismatch and the higher sound speed in water. An improvement in sound localization underwater will significantly enhance divers safety, the way divers perceive and appreciate the underwater environments. A system that augments and enhances the sound localization ability of humans underwater was built for this purpose. The system extracts directional cues from high frequency acoustic component of the received signal and reintroduce the cues in audio band to the diver that wears the system. The novelty of this approach is that it does not need any explicit information on the signals in advance to localize them. The system passes almost the entire signal band to its user with minimum relative distortion except the directional cue ti re-introduced. It is then up to the user to perceive, detect, and localize the sound.**

In this paper, we present the setup and results from an experiment that measures the localization performance of divers using the system. The experiment setup consists of a source transmitter that was randomly positioned in a contiguous, one-meter radius, semi-circular frame, and a blindfolded subject that attempts to localize the acoustic source. Both the headings of the transmitter and subject were digitally recorded and compared to gauge the localization performance. Experiments have been carried out across different signal to noise ratio and across different frequencies above 20kHz. The result from the experiment shows that a diver using the system was able to localize a source to within $\pm$15 degrees nearly 75% of the time. It is also observed that SNR does not significantly affect the localization performance within the range of SNR that we were testing. The subjects were able to localize acoustic source in a noisy marina environment with the system. The localization performance of the subjects seemed to improve as the subjects gained experience using the system over a few experiment sets. This suggests that the human brain adapts its perception ability and learns to use the new directional cues rather quickly.

I. Introduction

Human have only two audio receivers (left and right ear) but are able to accurately localize sound in three dimensional space. This is possible because the inner-ear, outer ear and all the related body parts modify the incoming sound and our brain extracts directional cues by processing the received sound [1]. These directional cues are heavily distorted if not lost once we are underwater, leaving our brain with inadequate cues to effectively perceive direction. This makes us very poor at localizing sounds underwater [2].

A pair of baffles were designed to restore some of these loses underwater in order to partially re-introduce the directional cues. This would therefore bring the directional audio perception that we are familiar with into the underwater world. The baffles introduce directional cues in the form of Inter-aural Intensity Differences (IID) with beam-patterns that mimics the Head-Related-Transfer-Functions (HRTF) in air. These directional cues are extracted from high frequency component of received acoustic signals and presented to subjects in audio frequency using a compression algorithm known as the Acoustic Bandwidth Compression (ABC) algorithm [3].

A set of swimming pool tests had been carried out in 2008 [4] to study the preliminary localization performance. In the experiment, five transmitters were deployed but only one of them was randomly selected to transmit a signal at any time. A diver subject was then tasked to predict and select which transmitter was active during the test. The transmitters were placed on a semi-circle with one meter radius from the subject about eye level but in different, discrete azimuth directions. The azimuth angular separation between the transmitters were then varied, allowing us to study the resolution of the localization. There was a possibility that the results from this experiment could have been psychologically affected, as the positions of the transmitters were discrete and visible to the subjects.

When an experiment was planned for open water trial, a new experimental setup has been designed to remove the limitation of the previous experiment. This paper describes the setup and results from this experiment.

As opposed to using a number of transmitters located at discrete positions, only one transmitter was required in the new setup. The transmitter was randomly positioned in a contiguous angular positions, and the subjects were tasked to localize it without visual reference. The heading measurements were no longer discrete: a compass was mounted at the transmitter to measure its angular position and another compass was mounted at the mask of the subject to measure his or her heading. The positions of the subjects were recorded throughout the experiment. These were used to calculate the subject's relative heading towards the source. This allowed the subject to rely only on his/her hearing to localize the sound with comfort. This was expected to remove any effect in the 2008 experiment caused by having a visual selection

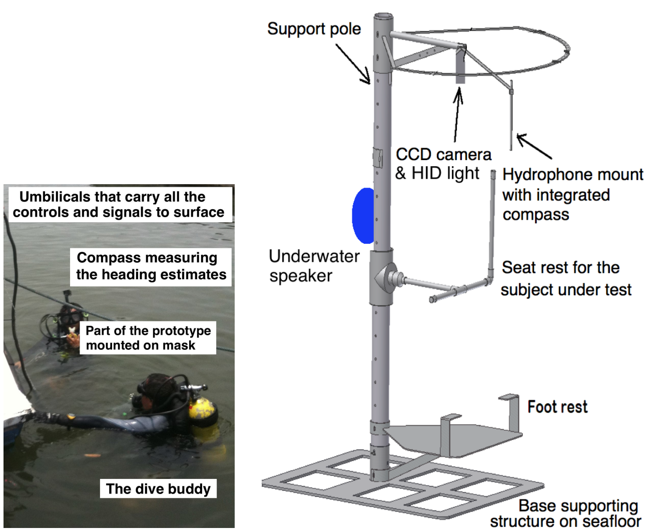
978-1-4244-4333-8/10/$25.00 ©2010 IEEE

of source transmitters and allowed greater resolution in the measurement.

II. Experiment Setup

As shown in Figure 1, the experimental setup consists of two functionally separate component. The first component is a prototype electronics that augments and introduce the sound localization of the divers, as shown in Figure 1(a). Figure 1(b) shows the second part of the setup that automates the experiment and tracks both the angular and translational displacements of the transmitter and subject simultaneously.

The prototype is integrated on the mask of the subject, coupling the subject's heading directly with the localization estimates. A reflective marker is placed at the prototype on top of the subject's head. An underwater camera is placed over the subject to measure the translation of the subject from the point of origin by locating the reflective marker. The underwater camera also serves as a secondary measurement unit of the subject's head orientation. One digital compass is placed on the transmitter mounting to measure the angular placement of the transmitter. A second digital compass is placed on the mounting of the receivers to measure the orientation of the prototype. This prototype is then fitted onto the subject's dive mask. Before the start of the experiment, both compasses were calibrated to correct for the magnetic influence of the environment. The data collection and processing were done on a desktop computer at the surface,



(a) Divers on the surface and the compass that measures the localization estimates (b) Experimental setup that automates the trial procedures

Fig. 1. Experiment setup

A. Experimental procedure

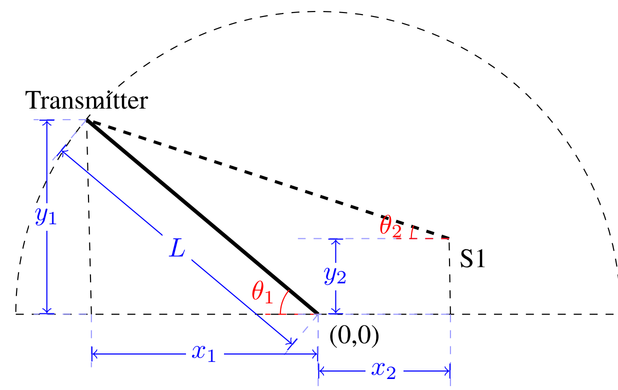
The experiment starts with a subject and a buddy underwater. An underwater speaker is placed near the buddy so that the buddy is notified by the surface controller at the start of the experiment. Each experiment consists of 25 individual tests. At the beginning of each tests, the buddy is to move the transmitter to any arbitrary angular position measured by compass A. During the test, the transmitter emits a stream of acoustic test signals.

The prototype continuously acquires any acoustic signal underwater. It compresses the acquired signal into audible sound, and plays back to the subject via an underwater headphone.

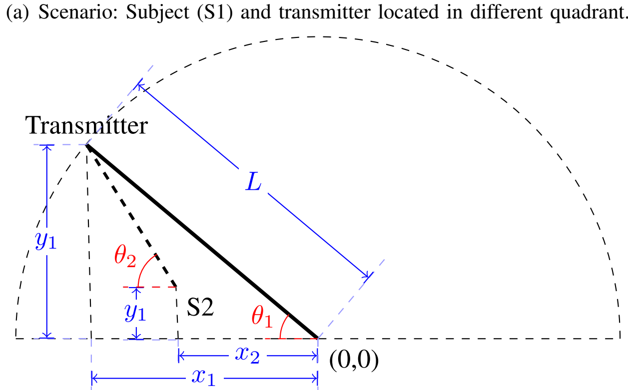
The subject is then given a certain amount of time to localize the sound source. The subject has a button which he/she presses to stamp the reading of all the compasses and timestamp the acoustic and video recordings. The process repeats itself for 25 times.

B. Compensating for translational offset

The experiment setup accounts for the displacement of the subject from the point of origin and calculates the subject's orientation with respect to the angular position of the transmitter. Figure 2 shows the basic geometry of the calculation.



(a) Scenario: Subject (S1) and transmitter located in different quadrant.



(b) Scenario: Subject (S2) and transmitter located in same quadrant.

Fig. 2. Aligning the subject's heading to the transmitter. With θ_1, x_2, y_2 measured in the experiment, θ_2 can be calculated.

There are two scenarios of the geometry but same mathematical concept are employed for both scenarios. In this figure, θ_1 and x_1,y_1 represents the angular displacement and the linear displacements of the transmitter from the origin. x_1 and y_1 is derived from L and θ_1. θ_1 is given by the compass that measures the transmitter heading. L is the distance of the transmitter from the origin, the point the compass pivots on. As the transmitter is mounted on a aluminum extension which can be swiveled a half circle, L is always constant. θ_2 is the angle of the transmitter with respect to the subject's position, geometrically derived from x_1, y_1, x_2, and y_2.

The displacements of x_2 and y_2 are measured through the image processing of the video recorded by underwater camera.

θ_2 is then compared to θ_3, the reading from the compass mount on the subject, as performance criteria.

It's worth noting that some amount of error could be introduced by external factors such as subject behavior and camera placement. From the experiment images, the displacement of the tracker from the origin is in the range of 20-30cm. A roll and pitch motion might be perceived as a translation up to $\pm$3cm in the image recorded by the camera. As the angular measurement is derived from translational displacement, these deviation will be translated into angular errors of 3-5 degree.

III. Swimming pool experiments

A number of swimming pool experiments were carried out in the beginning. They served as familiarization runs to prepare the subjects for the open water experiment which is a more complex environment. It was also set up as a quick test of the hypothesis of possible psychological bias in the 2008 experiment due to the presence of visual reference.

The experiment was carried out with a fixed source level, same as the experiment carried out in 2008. Similarly, each tests were also limited to 35 seconds. During the experiment, a subject with scuba gear was placed in the experiment setup. The localization unit was mounted at the mask and the subject was required to run the experiment with eyes closed. A swimmer was deployed in the pool to randomly position the source transmitter in between tests. The subject would then localize the source by facing the source. The headings of both transmitter and subject were recorded, along with the acoustic signal at the time.

A total of 125 tests were carried out, spanning 5 experiments. The raw compass readings of the source transmitter were post-processed and aligned to the body positions of the subjects instead of the origin as described in earlier section. These were then directly compared with the source heading estimates of the subject recorded by the compass. The error of the source angular position estimates were then tabulated and their histogram is shown in Figure 3.

It is observed that the subjects were able to estimate source direction with error of less than $\pm$22 degree about 75% of the time. Results from 2008 experiment showed success localization probability of more than 82% at spacing equivalent to $\pm$22.5 degree resolution (spacing of 45 degree). Whilst a 75% success detection rate yields equivalent of $\pm$12.5 degree resolution (at 25 degree spacing) [4]. This seem to support the hypothesis that the experiment setup with visual reference might have over-estimated the performance.

IV. Open water experiments

A set open water experiments were carried out following the swimming pool tests. The site of the experiment has much greater environmental disturbance as compared to swimming

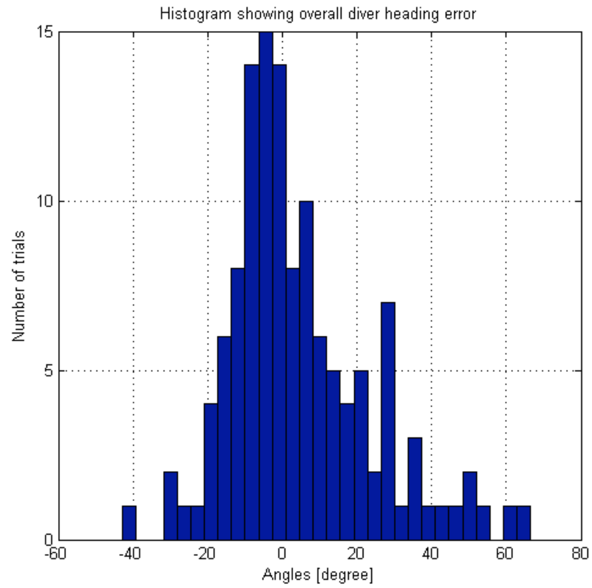


Fig. 3. Histogram of errors of heading estimates ($\theta_3 - \theta_2$) from swimming pool experiment from all subjects

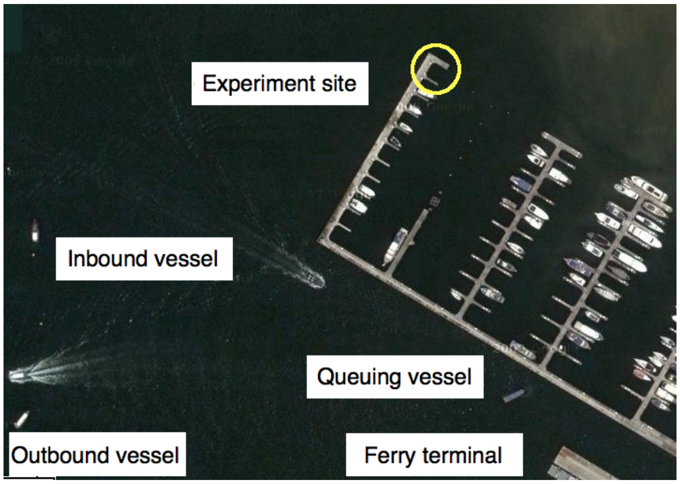
pool environment, which represents a more realistic operating environment. We started the experiment with fixed frequency pulses and fixed source level, similar to the swimming pool test. Once the subjects were familiar with the environment and had obtained similar performance as swimming pool trials, we moved to a more elaborate test scenarios involving multiple frequencies and SNR levels.

The same experimental setup was used with a couple of upgrades that were essential to operate in poor visibility. These include an underwater HID lamp to locate a reflective tracker on subject, and an underwater speaker to coordinate the experiment acoustically from the surface. The main difference in the experiment setup was that the subjects were allowed as much time as they needed to localize the sound, instead of limiting it to 35 seconds. Two divers were deployed in each experiment, one as subject and the other as dive buddy. The dive buddy was to support the subject during setup, communicate with surface as when needed, and randomizing the source location between each tests. Instructions were relayed verbally to the divers through the underwater speaker, while diver replied with hand signals hold near the camera.

Twenty-six sets of experiments were performed, which yielded eleven sets of useful data. The discarded data sets were due to poor visibility that prohibited the data to be effectively processed (11 sets), a compass failure (1 set), and operational errors (3 sets). The rest of discussions in this paper is based on results from these open water experiments.

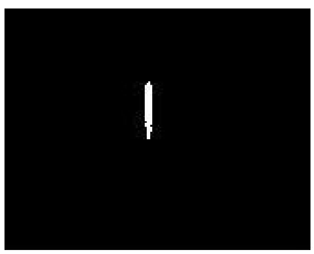
A. Trial site

The site of the open water experiment is a marina. It is located less than 100m meters away from a busy ferry terminal, where there is a constant small craft activities, either inbound or outbound. The experiment was setup at a seaward pontoon with about 6m water depth, see Figure 4(a).



(a) Experiment site. Note the busy ferry terminal south to the site

(b) View of diver from camera in good condition



(c) Extracted diver position after image processing

Fig. 4. Condition of trial site.

The visibility was normally less than 0.5m near the seafloor, or worse when the silt or mud was stirred up from the bottom when divers were settling in the setup. A high power underwater High Intensity Discharge (HID) lamp was used to illuminate the reflective marker placed on the subject located less than 0.5m away from a camera. This produces a region with distinctively higher illumination on the marker compared to the background. Image processing is then carried out to threshold the pixel value and isolate the marker from the backgrund and produce a black and white image, see Figures 4(b) and 4(c). The coordinate of the centroid of the isolated region is then measured to locate the subject's position.

The subject's position is crucial to compare the heading of the subject to the actual location of the source.

B. Notion of SNR definition used

During the experiment, one of the objectives was to investigate the performance for different SNR. Assuming that divers are able to extract the signal of interest and filter the ambient noise in the perception process, then out of band noise spectrum will have negligible effect to the perceived SNR. Hence, we calculated the SNR by only taking the time section when the signal is present and spectral band of the signal. For example, for a signal with bandwidth b kHz, and center frequency f_c, the signal power and noise power is the sum of their respective spectral estimates across b kHz centered around f_c. Hence the band limited SNR,

$$SNR_{BL} = \frac{\int_{f_c-b/2}^{f_c+b/2} s(f)df}{\int_{f_c-b/2}^{f_c+b/2} n(f)df}$$

Where,
$s(f)$ = spectral density of signal of interest
$n(f)$ = spectral density of noise

$n(f)$ is interpolated from the nearest lower and upper spectral density estimates immediately outside the frequency band of the test signal.

C. The challenge of rapidly fluctuating ambient noise at the trial site

Due to the presence of a variety of active, small, water crafts in proximity, the ambient noise in the area fluctuates greatly and rapidly. this is mainly due to the engine noises that consist of rumblings of various lengths as well as transients. Preliminary test with constant source level resulted in SNR variation of 20dB. This suggested that ambient noise level in the area can vary within 20dB during the experiment. This made the generation of source levels at desired SNRs a non-trivial task.

In order to tackle this issue, the source level was adapted to the ambient noise at the time of the test. The ambient noise was sampled and used to calculate the appropriate source level based on the required SNR at the beginning of each tests. Although this does not completely eliminate the SNR fluctuation, it reduces the fluctuation level from 20dB to about 6dB more than 75% of the time. This allowed us to operate the experiment near the desired SNR levels. The actual SNR of each tests was calculated from the acoustic recording and used in the analysis of the results.

The following sections describes a number of observations from the experiment results.

V. The Learning Capability of Subjects Using the Augmented Localization System

A number of constant source level experiments have been carried out. The source level was chosen such that the subjects can easily detect the signals. This allowed the subjects to familiarize with the experimental flow and the new environment. It also allowed us to concentrate on solving the technical issues during the trial. Four sets of successful experiments have been carried out, out of which 56 data points have sufficient image quality to be processed.

The top plot in Figure 5 shows the error distribution in estimating the source direction against SNR levels. The SNR level of each tests was calculated from the recorded acoustic data. The majority of the tests seemed to yield SNR within 45 to 50 dB except 2 tests points that lean towards lower SNR.

About 80% of the 56 tests have maximum localization error of ± 50 degrees.

There are a number of large localization errors across different SNR level. This led us to suspect the possibility of the presence of a learning curve associated to the operation of the new setup and/or the new site. The error distribution was then plotted with decreasing color intensity in chronological order, i.e. the darkest points were the data from the first test of the first experiments and the lightest points were from the last test of last experiment. It indeed showed a trend of reduction in localization errors over time.

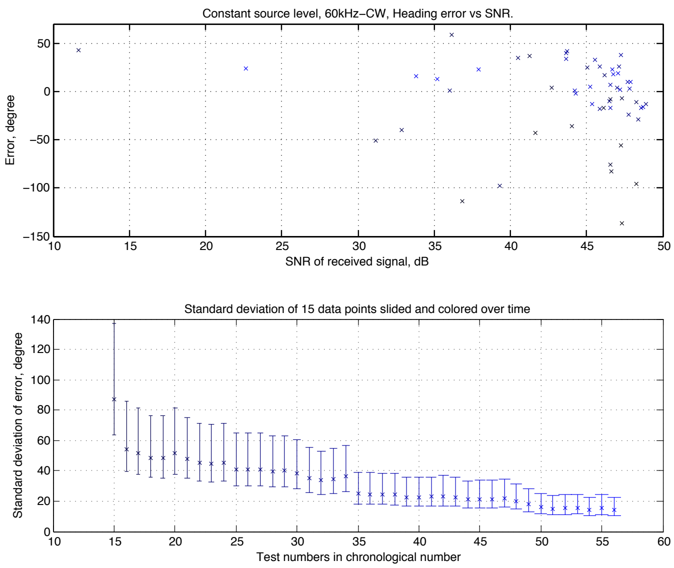


Fig. 5. Results of constant source level experiment, with the color darken over time.

We further examined the change in localization performance over time by observing how the standard deviation of the error varies over time. The standard deviation of the error at time t, s_t, was calculated over a sliding window of n tests in sequence.

$$s_t = \left(\frac{1}{n-1} \sum_{i=t-n+1}^{t} \left(\theta_i - \bar{\theta}_t\right)^2 \right)^{\frac{1}{2}}, \ \forall t = n, n{+}1, ..., N{-}1, N$$

where, $\bar{\theta}_t$ = mean of the error estimates for $\{\theta_{t-n+1}, ..., \theta_t\}$. and, N = the number of tests

n was chosen to be 15, a number that is large enough to filter fluctuations in the the estimates yet small enough to shows the trend correctly. N was taken as 56, the total number of useable data points from these four experiments.

The lower plot in Figure 5 shows the standard deviation against chronological test numbers. The error bars in the figure show the 95% confidentce interval.

It is observed that standard deviation of errors in localization estimates reduces over time. This suggest that the subjects were able to rapidly learn the system and the environment to be proficient with the system. In this case, they were able to reduce the standard deviation of localization error from about 50 degrees to about 15 degrees in just a few trials.

Although a subject that was new to a new setup, in a new environment will have poorer localization performance, the localization error were typically less than 55 degrees. The result also suggests that the standard deviation of error reduces and stabilizes around 15 degrees. Lastly, it also appears that SNR of the signal has little effect to the localization performance.

VI. The Resilience of Localization Performance at Different SNR Levels

We next tested the effects of the SNR levels on localization performance. A total of 69 valid data points were obtained, and the localization estimates were corrected for the translational offset of the subject's positions and the errors were calculated. A Lilliefors test with 5% significant level was carried out to investigate if the error distribution was normal. The test did not reject the null hypothesis that observed data is normal distribution with p-value around 0.118. Nevertheless there were four localization errors that deviated significantly more than the rest. These data points were regarded as outliers and removed from the rest of the analysis so that we get a better consistency in the statistical analysis that assumes normal distribution. A repeated Lilliefors test resulted in a better normal fit with p-value of 0.379. Figure 6 shows the localization error distribution fits a normal distribution well with a mean of -13.4 degrees and standard deviation of 14.4 degrees.

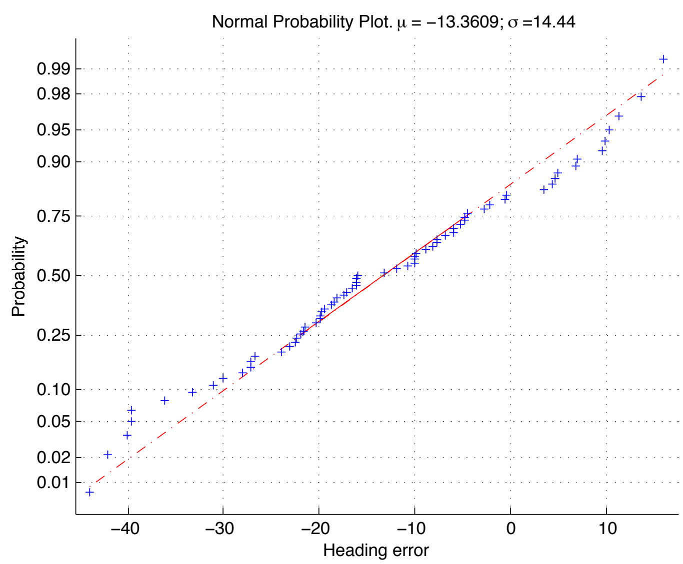


Fig. 6. Normal fit of the data.

The errors of heading estimates were grouped into bins of SNR with 5dB and 10dB increments and the standard deviations of each bins were calculated. Figure 7 shows errors in heading estimates and the standard deviation of the bins against different SNR A 95% confidence interval error bar was calculated and plotted along with each of the standard deviation of error estimates. Note that standard deviation from bin 10-15dB and bin 30-35dB have much larger error bar because their number of data sets are too small to give good estimate. The plot shows no evidence that the standard deviations of the errors in heading estimation are affected by the changes of SNR levels. This suggests that the subjects were able to estimate direction at a similarly accuracy from 5dB to 50 dB SNR.

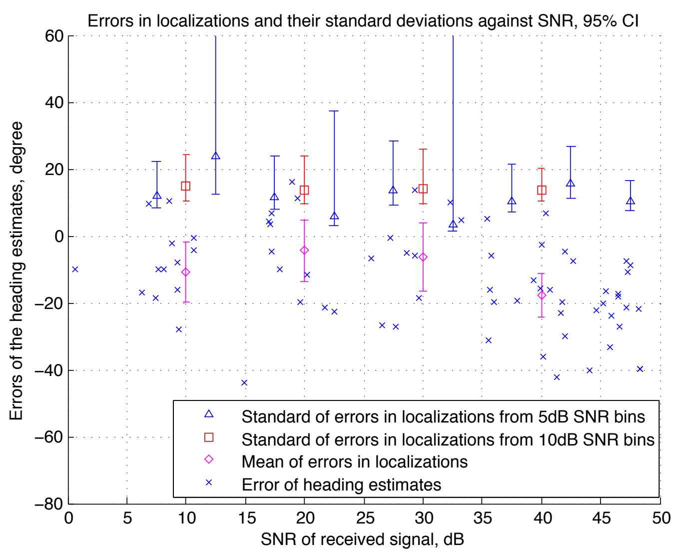


Fig. 7. Error of heading estimates versus SNR. The diamond-markers indicates the mean estimates, the square and triangle-markers indicates standard of error.

Figure 7 also shows the mean of the error estimates over SNR bins of 5-15dB, 15-25dB, 25-35dB, and 5-45dB along with their 95% confidence intervals. The overall mean estimates of -13.4 degrees fit well within the 95% confidence interval error bars across all of the bins. This suggest that there is consistent offset in directional perception using the prototype. This offset in localization error mainly came from the offsets in the augmented perception which are both individual and system dependent. These can be removed by performing a calibration at the beginning of the experiment, and this has been demonstrated in the experiment carried out in 2008 [4]. The calibration procedure was not employed in these experiments in order to make full use of the experiment time for data collection. Instead of the mean of the error, the dispersion of the heading estimates are used as the indicator of how accurately one could localize the direction of the source. This is measured by the standard deviation of the localization errors.

With ambient noise level of frequency range that we planned to operate in and employing spreading and simple absorption model, the SNR range would be equivalent to physical distance of 1km to 3.7km. This is consistent with a subsequent open sea test, where a subject was able to localize the position of a source located nearly 1.3 km away with an error of 6 degrees.

This suggests that the subjects were able to localize a source as long as one can perceive it, even at very low SNR.

VII. Conclusion

We have presented an experiment carried out to test a prototype system that augments sound localization of divers in open water. The performance of the prototype system has been studied in a dynamic environment with strong, fluctuating acoustic disturbance. The results have drawn some interesting observations that are encouraging.

The subjects in the experiments were able to proficiency localize test signals well, even at very low SNR levels. The overall error of the localization estimates were measured to have standard deviation of nearly 15 degrees, i.e. we were able to accurately localize to ±15 degrees about 75% of the time. This has lower performance indication than the 2008 experiment result that estimated a 75% probability of success localization with ±12.5 degrees resolution. This is possibly due to the difference in experiment design, where the setup in 2008 experiment might have over-estimates of localization performance. We also found that SNR level has little effect on the performance. Localization is highly feasible as long as the subjects were able to reasonably hear the signal. Lastly, it is also clearly observed that the subjects were able to learn the system and improve the localization perception over time.

Knowing that the directional perception is possible underwater using the system, next step of work would be implementing a diver wearable portable electronics package. This would allow us to study how well the augmented directional perception work with natural underwater sounds. Exploration of visual and audio cross-model localization would be an interesting topic to investigate.

Acknowledgment

The authors would like to thank the Royal Singapore Yacht Club for the support in providing a dive site for this research. We would also like to thank Alan Low and other volunteers that have helped out this research.

References

[1] C.I. Cheng and G.H. Wakefield, *I*ntroduction to Head-Related Transfer Functions (HRTF's): Representations of HRTF's in Time, Frequency, and Space (Invited Paper), Proceedings of the Audio Engineering Society 107th Convention, New York: 1999.

[2] Harry Hollien, *U*nderwater sound localization in humans, J. Acoust. Soc. Am. 53, 1288 (1973)

[3] T. B. Koay, J. Potter, M. Chitre, S. Ruiz, and E. Delory, *A* compact real-time acoustic bandwidth compression system for real-time monitoring of ultrasound, in OCEANS '04. MTTS/IEEE TECHNO-OCEAN '04, vol. 4, pp. 2323-2329, November 2004.

[4] T. B. Koay, J. Tan, S. P. Tan, S. K. Yeo, H. Tay, M. Chitre, and J. Potter, *E*nabling humans to hear the direction of sounds underwater - Experiments and preliminary results, in OCEANS 2008, pp. 1-9, September 2008.