An Underwater Wearable Computer for Two Way Human-Dolphin Communication Experimentation

Daniel Kohlsdorf[1], **Scott Gilliland**[1], **Peter Presti**[1], **Thad Starner**[1], **and Denise Herzing**[2]
[1]Georgia Institute of Technology, Atlanta, GA and [2]Wild Dolphin Project, Jupiter, FL
(dkohlsdorf,scott.gilliland,thad)@cc.gatech.edu, peter.presti@imtc.gatech.edu,
wdpdenise@wilddolphinproject.org

ABSTRACT

Research in dolphin cognition and communication in the wild is still a challenging task for marine biologists. Most problems arise from the uncontrolled nature of field studies and the challenges of building suitable underwater research equipment. We present a novel underwater wearable computer enabling researchers to engage in an audio-based interaction between humans and dolphins. The design requirements are based on a research protocol developed by a team of marine biologists associated with the Wild Dolphin Project.

Author Keywords

underwater wearable computers, animal vocalization, hydrophones, dolphins

ACM Classification Keywords

H.5.2 Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Marine biologists of the Wild Dolphin Project are interested in observing social structures and communication in the wild. The question for marine biologists is will wild dolphins not only mimic computer generated whistles but also interact using theses whistles? In an experimental protocol designed by the Wild Dolphin Project, one person or dolphin sends an audible request (a whistle) for a certain object, for example a scarf. A researcher possessing the requested object will then pass the object to the person or dolphin sending the request. We introduce an underwater wearable computer supporting this protocol to engage in an audio-based human-dolphin interaction. The system enables a diver to play a predefined set of artificial whistles using a keyboard. Furthermore, the system can recognize these whistles using input hydrophones and attempts to estimate the source's direction. We discuss the hardware design as well as the customized software. Furthermore, we discuss our performance measurements from three field deployments in the open ocean off the coast of the Bahamas and several pool tests.

RELATED WORK

Previous research has shown that captive dolphins have the ability to mimic whistles with high accuracy [4]. However, the researchers could not respond or engage in an interaction because they did not have an online recognition and playback system. Working with wild dolphins, Herzing et al. [3] used an underwater keyboard with symbols and artificial whistles associated with certain objects. They observed that human use of the system encouraged the dolphins to participate.

Underwater wearable computers are widely used in scuba diving for measuring pressure, depth and temperature, and controlling the mix of gases from the tank in rebreathers. More relevant to this work is the OTS Dive Buddy [1] which allows teams of divers to communicate wirelessly using ultrasound while underwater using a microphone and headphone integrated into their snorkel or mask.

Our whistle recognizer is based on the algorithm used in the Shazaam application [5] for mobile phones. Shazaam can recognize songs with high accuracy in a noisy environment using a 30 second snippet.

A BINAURAL UNDERWATER WEARABLE COMPUTER

Designing a system for operation in a saltwater environment presents a number of challenges. The most immediate concern is water infiltration. We created a case comprised of a custom machined aluminum housing with a clear Lexan back plate. An O-ring compressed by 14 socket head bolts seals the box. External devices are attached to the box with IP67 sealed multi-pin connectors or compression fittings on cables. The system is shown in Figure 1.

A second Lexan plate bolted to the front of the system holds two hydrophones suspended with rubber shock mounts. The mounting isolates the hydrophones from each other and vibrations in the box. This plate also supports a custom piezo speaker built from a piezo element cast into a urethane rubber. Off-the-shelf Audio Bone^{TM} headphones provide audio feedback to the user. A custom keyboard with display is mounted above the speaker plate. The interface consists of 12 sealed tactile buttons, an edge-mounted 4 character display, and a communications cable for I2C signals back to the host. Located inside the box is a Fit PC and an attached iMic USB audio interface that captures a stereo signal from the input

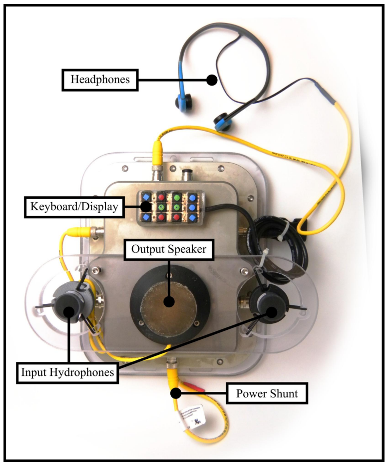

Figure 1. The underwater wearable computer. Two input hydrophones are mounted at the sides, while the output hydrophone is mounted on the front along with the keyboard.

hydrophones. A class-D audio amplifier drives the speaker and headphones. A 20x2 LCD provides debugging information visible through the Lexan cover. Finally, a 11.1V 6Ah lithium polymer battery pack powers the system. The battery is charged through a connector at the base of the unit, and the system is switched on by installing a watertight shunt on the charge port. The system operates for 8-12 hours on a full charge and is typically charged overnight at a rate of 1Ah for 7 hours. Due to the high cost and difficulty driving underwater piezo units, a custom underwater speaker was constructed. The speaker was assembled from a PUI Audio AT-5210-TT-R piezo transducer potted with a Shore 44A hardness polyurethane rubber in a 3D printed ABS housing.

WHISTLE RECOGNITION

We designed the artificial whistles to be easily reproducable by a dolphin and easy to recognize by our software. We added a start marker tone at 5 kHz and a stop marker tone at 7 kHz. The resulting whistles are single frequency sweeps between $\sim$3 kHz and $\sim$8 kHz. The recognition is implemented in three different components. A start and stop marker detector, a whistle classifier, and a sound source localizer (ssl). During recognition we maintain a $\sim$0.07 second sliding window for both channels with a $\sim$0.05 second overlap. For each of these windows we extract the log-abs magnitude of frequencies between $3kHz$ and $8kHz$ using Goertzel's algorithm [2]. The start and stop detector triggers if the dominant frequency is in the $5kHz$ or the $7kHz$ band. If the start detector triggers, we classify the following sliding windows between the start and stop marker using a recognizer based on the Shazaam [5] algorithms with slight modifications. The Shazam algorithm searches significant points in frequency in the spectrogram allowing only exact matches. In contrast, our algorithm uses a threshold based classifier on the distance between stored points and points from incoming audio during live recognition. In that way we allow points to vary in frequency and time allowing additional robustness. Once we recognize a whistle, we compute a source location (to the left, center, or right of the researcher) using only the last sliding window from each channel. Currently, we use the general cross correlation with phase transition for localization.

Testing

Since the whistles are synthetic and consist of single frequency sweeps, the problem is not the recognition itself but distinguishing between the whistles and other underwater sounds such as waves and boats. The team conducted an initial experiment using two researchers wearing the system. The researchers played back the sounds several times at four, seven, ten, thirteen and eighteen meters in a diving pool. The recognizer correctly identified the whistle 98% of the time even when we created splash noise in the water. One false negative occurred, resulting from an incorrect audio gain setting. There were no false positives observed. Thus, the resulting precision is 1 and the recall is 0.98. The start and stop detector prevent false triggering since ambient non-artificial sounds do not occur as a loud standing sine wave. The Shazaam algorithm proved robust to noise as well. The team then tested sound source location. In the pool test we evaluated performance at a seven meter distance with angles ranging from -90 to 90 degrees in 45 degree intervals. While we had the conservative goal of detecting if a sound was made to the left, center, or right of the user, the system performed poorly.

FUTURE WORK

We are improving the localization system before ocean experiments this summer. Using the data collected from the pool test, we are re-estimating the thresholds used for directionality and will test and confirm operation in a pool. We will also explore other avenues such as comparing the relative loudness of the signals hitting both input hydrophones and moving the hydrophones side by side so as to capture phase differences in the signal to determine signal location. Before field trials, we will test in the open ocean as the ocean has significantly higher levels of background noise.

CONCLUSION

We developed an underwater wearable computer enabling research into two way human-dolphin communication. The system is able to recognize and produce computer generated whistles. In a controlled pool deployment we found that the system is reliable under noise at a distance of at least 18 meters under water. In the future, we will deploy the system in the Bahamas, gathering more data as well as using the system for field trials with wild dolphins.

REFERENCES

1. Diving buddy. `ots.mwrc.net/en/product.php?product_id=45400`.
2. Goertzel, G. An algorithm for the evaluation of finite trigonometric series. *American Mathematical Monthly* (1985).
3. Herzing, D., Delfour, F., and Pack, A. Responses of human-habituated wild atlantic spotted dolphins to play behaviors using a two-way human/dolphin interface. *J. of Comparative Psychology* (2012).
4. Richards, D., Wolz, J., and Herman, L. Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, tursiops truncatus. *J. of Comparative Psychology* (1984).
5. Wang, A. An industrial-strength audio search algorithm. In *Proceedings of the 4 th Intl. Conf. on Music Information Retrieval* (2003).