Sing Me a Song! - Trials and tribulations of a first attempt to record humpback whale mother-calf pairs in the DR with a new version of a synchronized video/audio array.

Hoffmann-Kuhnt, Matthias
tmsmh@nus.edu.sg

Ho, Abel
abelho@nus.edu.sg

Vishnu, Hari
tmshv@nus.edu.sg

Acoustic Research Laboratory, Tropical Marine Science Institute, National University of Singapore

Abstract

Based on our previous recoding system we developed an adaptable advanced video/audio recording device that consisted of a 6k resolution camera (Z-cam E2S6) with an ultrawide angle lens (170 degree FOV) , 4 hydrophones (Reson 4013 with custom in-line amplifiers) a data acquisition card with a sampling rates of 400 kSamples/second/channel, custom second-stage amplifiers and custom power management. The system was housed in a 6-inch diameter custom-built aluminium tube with a Nauticam fisheye dome port in the front plate and control switches and LEDs in the back plate. Attached to the housing were 3 aluminium arms extending out to form an equilateral triangle with three of the hydrophones in each of the corners and the camera in the center of that triangle. A fourth hydrophone was mounted with a 30 cm offset of the plane of the other hydrophones to eliminate a possible front/back ambiguity for localization of sources. Different length arms could be attached to the housing to form smaller and larger arrays – this allowed us to make a large enough array (120cm hydrophone separation) to record and identify low frequency sources (humpback whale mother/calf pairs). This system was deployed for the first time in the waters of the Dominican Republic (Naviad Bank and Silver Bank) in March of 2021. We report here the first attempts of deployment and recording with this system and the analysis of the results.

Experimental Setup

A previous version of the recorder was based on an underwater video housing that contained all the electronics (DAQ board, amplifiers, IP-camera, power regulation and batteries) but the hydrophones with built-in preamps were mounted externally in a fixed equilateral triangle array on the same plane as the camera. This older version had a separation of 60 cm between the three hydrophones and that setup was working well for the recordings of higher frequencies (i.e. smaller cetacean species such as bottlenose dolphins and spotted dolphins), but was not sufficient to record low-frequency vocalizations of larger whales such as humpbacks.

In the fall of 2020, the BBC approached us to learn more about the recording capabilities of this device and ask if it was possible to apply the method to humpback whales' mother-calf pairs – and we said: "In principle yes" - but modifications would be necessary. We then started modifying the older existing device to achieve an array size of 120 cm separation between hydrophones to adapt the system to a lower frequency range while simultaneously applying for a National Geographic Society Tech Grant to build a completely new system that would improve upon the existing system in several areas:

1) The previously-used IP camera that had a resolution of 1080 and a diagonal FOV of about 70 degrees would be replaced with professional video camera – a Z-Cam S6 -E2 with a resolution of 6K accompanied by a Rokinon 8-mm fisheye lens with a diagonal FOV of approximately 170 degrees.

2) The hydrophones would be mounted in such a way that the separation between the three hydrophones would be 120 cm – large enough to adjust to the lower frequencies but still maneuverable as a whole system by a swimmer/snorkeler.

3) A fourth hydrophone would be added to eliminate the front/back ambiguity of sound sources in respect to the plane of the other three hydrophones.

4) Improved power management and tethered Ethernet connection would also be added that would allow us to deploy the system on a tether from an small boat and still be able to monitor the orientation and visual capture of the device.

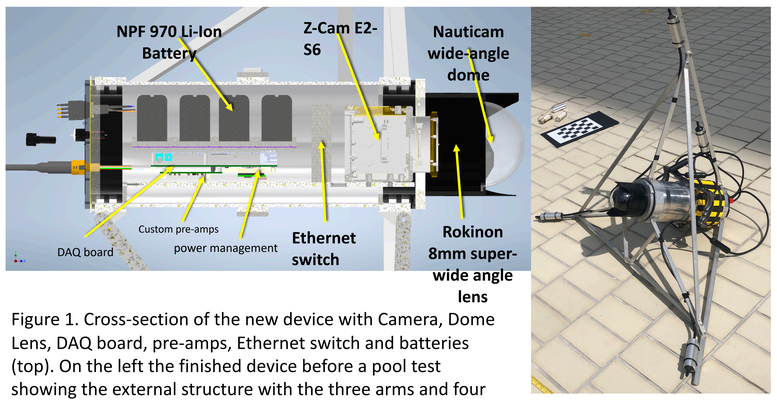

Figure 1. Cross-section of the new device with Camera, Dome Lens, DAQ board, pre-amps, Ethernet switch and batteries (top). On the left the finished device before a pool test showing the external structure with the three arms and four hydrophones.

Deployment

The system was developed, put together and tested in Singapore before being deployed in the Dominican Republic in March of 2021. At that point in time, we had been able to put together a rudimentary system – that still lacked the tether and remote viewing ability. The system deployment was from the OceanXplorer II (https://oceanx.org/oceanxplorer)– a vessel solely dedicated to ocean exploration and research – that was, at the time, conduction research in the area around Silver Bank and Navidad Bank off the coast of the Dominican Republic.

Unfortunately, the time to acquire data was very limited as the research vessel was in the area of the Dominican Republic only for about 7 days. Additionally, the weather forecast predicted a storm for 3 of these days and the ship had to seek shelter in Samana Bay. Thus, realistically left with only 2 days of possible recordings, we did our best to record any whales that were in the area – preferably mother-calf pairs (MCE).

A smaller tender boat (the FRC) was initially used to search for whales and to deploy the equipment – but it turned out that the engines (while very powerful) where too loud and when we approached a mother-calf pair they would simply vanish. We then switched to a smaller inflatable zodiac that had quieter engines.

Figure 2. the OceanXplorer from afar, the FRC (top right) and the zodiacs used (bottom right)

Eventually, we had very few encounters in total that we were able to process:

- – a mother-calf-escort group where we had a very low frequency groan in the recording and where the SNR was very low – so the localization remained inconclusive.
- a singer that was very audible but not visible on the camera - - we could detect his location only acoustically
- - And finally a mother calf-escort group that was visible on the video fairly close up – but did not vocalize at all.

Data processing

The sound locations were estimated using a time-domain Bartlett beamformer, scanning in a 2D elevation-azimuthal grid of -80 to 80 degrees (in both directions). All 4 hydrophones were used. The time-domain beamforming allowed us to utilize the entire bandwidth of recorded sound available, thus allowing us to mitigate grating lobes in the output when enough broadband information was available. The time-differences of arrival at the different hydrophones gave information on whether the sounds were coming from the front or the back of the array (and camera), and if they were coming from the back of the camera, they were not registered or shown on the display. The results of the beamforming were then overlayed onto the the individual video frames that corresponded to the acoustic recordings. A 3-dB contour line was plotted at the location of the source to indicate the highest energy (see figure 3 ad 4). Next to the video frame a waterfall spectrogram was plotted that showed the frequency content of the current sound and below the frame a time series of the sounds was plotted to show the current location of the analysis window relative to the total recording.

Figure 3: the MCE group with a possible localization of a groan (low SNR)

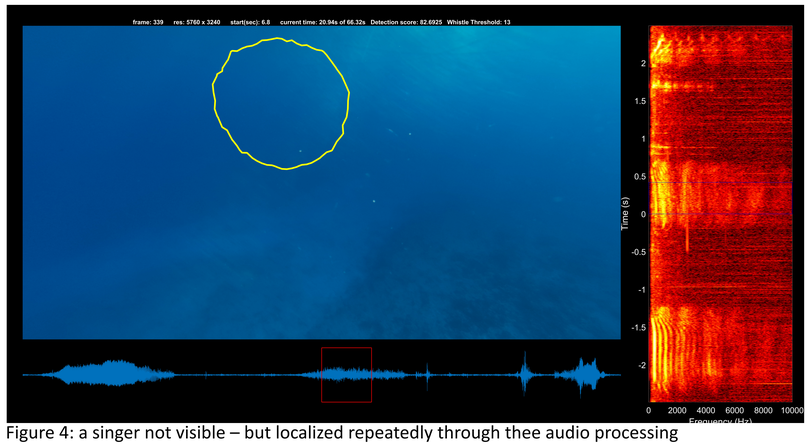

Figure 4: a singer not visible – but localized repeatedly through thee audio processing

Conclusion

While obtaining repeated good recordings from MC-pairs remained elusive in the very short time where we had access to to the whales, we were nevertheless able to demonstrate that the method applied here is sound (pun intended! 😜) and that over time with more and more recordings, localization of sounds produced by either the calf or the mother can definitively be identified. It is not surprising that mother-calf vocalizations are very quiet and rare (whispering), since the the animals do not want to advertise their presence to potential predators.

Furthermore, the equipment and processing technique can applied to a variety of sound-producing species over a broad frequency range – if the array size can be adjusted to the the bandwidth in question. Higher frequencies would call for a smaller array size while low-frequency vocalizations would need a larger array. Since the "arms" on which the hydrophones are mounted can be changed and adjusted different recording frequency ranges are now possible with the same equipment.

Acknowledgement

We would like to thank the BBC for their support with the development of the new equipment as well as the travel support to the Dominican Republic, the crew and staff of the OceanXplorer for providing a fantastic platform to work from and for going out of their way to get us to the elusive whales. And lastly to the National Geographic Society for funding this endeavor to build a complete new system from scratch in a very short time.

References

- Matthias Hoffmann-Kuhnt., Ho, A., Herzing, D., & Chitre, M. (2018). Whistling By – Issues with Identifying Moving Vocalising Dolphins. In 8th International Workshop on Detection, Classification, Localization, and Density Estimation of marine mammals using passive acoustics. France.
- Hoffman-Kuhnt, M., & Herzing, D. (n.d.). Whose Line Sound is it Anyway? Identifying the Vocalizer on Underwater Video by Localizing with a Hydrophone Array. Animal Behavior and Cognition, 3(4), 288-298.