Automated Video Surveillance for the Study of Marine Mammal Behavior and Cognition

Jeremy Karnowski[1,*], Christine Johnson[1], & Edwin Hutchins[1]

[1]University of California, San Diego
*Corresponding author (Email: jkarnows@cogsci.ucsd.edu)

Abstract - Systems for detecting and tracking social marine mammals, including dolphins, can provide data to help explain their social dynamics, predict their behavior, and measure the impact of human interference. Data collected from video surveillance methods can be consistently and systematically sampled for studies of behavior, and frame-by-frame analyses can uncover insights impossible to observe from real-time, freely occurring natural behavior. Advances in boat-based, aerial, and underwater recording platforms provide opportunities to document the behavior of marine mammals and create massive datasets. The use of human experts to detect, track, identify individuals, and recognize activity in video demands significant time and financial investment. This paper examines automated methods designed to analyze large video corpora containing marine mammals. While research is converging on best solutions for some automated tasks, particularly detection and classification, many research domains are ripe for exploration.

Keywords – Dolphins, Video data, Automated detection, Unmanned aerial vehicles, Underwater surveillance

Systems that detect and track social marine mammals provide data that can be crucial for understanding their social dynamics. Using these data, researchers can uncover regularities in long term behavior and build models to predict social interactions. Additionally, recording the position and movements of these mammals during social interactions is essential if researchers wish to establish more nuanced relationships between behavioral contexts and the vocalizations that occur within them. As recording technology becomes cheaper and more portable, researchers are collecting more data than ever before. Just as it was recognized that understanding animal communication systems would require advances in analytic tools for ever increasing datasets (McCowan, Hanser, & Doyle, 1999), it is now apparent that in order to adequately address and make effective use of this new wealth of video data, we must develop ways to automate the processing of it.

One possible non-video solution would be to attach devices, such as Digital Acoustic Recording Tags (DTAGs) (Johnson & Tyack, 2003), to every animal in a group in order to obtain GPS coordinates, depth, and movements for all interactions (Tyson, Friedlaender, Ware, Stimpert, & Nowacek, 2012). Although this technique gives very accurate location data, the limited availability and high price tag of DTAGs make it prohibitively expensive to tag many animals. Even with multiple tags, it would take considerable effort to attach a device to every animal in a study group and ensure that they all remain attached for an overlapping duration. In addition, dolphins will modify their swimming behavior when these tags are attached (van der Hoop et al., 2014) indicating that the tagging process is intrusive and may

interfere with group dynamics. Acoustic tags, or the use of hydrophone arrays, can help identify the position of animals over time through triangulation, but unless the animals are constantly vocalizing, they cannot be continually tracked.

Investigating marine mammal behavior and cognition requires a systematic and consistently applied sampling regime in order to produce unbiased data (Nowacek, Tyack, & Wells, 2001). Video surveillance constitutes an inexpensive and viable option for accomplishing this. Video provides the raw data - the physical record of events - from which conclusions can be drawn. Video provides an alternative to deciding on a schema for behavior ahead of time and allows for the re-sampling of behavior, the scoring of new features, and the reevaluation of activity. The ability to sample and retain hundreds or thousands of interactions between animals provides a substrate for understanding animal relationships at multiple levels. Researchers, after studying how relationships develop, can re-investigate how they code for and interpret the interactions underlying those relationships. Having a digital record supports the ability to re-evaluate the same data given new hypotheses (Forster, 2002). Finally, video provides us with the ability to slow the viewing rate to reveal previously undetectable behaviors. Frame-by-frame analyses can provide the minute details that underlie and help explain behavioral sequences and their dynamics (see Herzing & Johnson, 2015; Johnson, 2001, 2010).

Video data collected in these studies require skilled annotators to (1) detect and localize animals, (2) track individual animals, (3) identify individual animals, and (4) classify behaviors/activity. Each of these tasks is incredibly time consuming and prone to errors. For example, it can take, on average, one hour to analyze and annotate one minute of video when conducting such micro-analysis, and this grows exponentially more time consuming as one increases the size of the video network.

The following sections will cover the variety of video collection methods currently being employed in research as well as the automated methods that have helped tackle tasks typically designated to trained coders.

Current Research

Boat and Shoreline Based Methods

Data collection. Traditional boat based observation enables researchers to apply their expertise to identify species, discriminate individuals, and recognize individual or group behaviors at the ocean's surface. With enough training, researchers can become quite skilled at these classification tasks. Researchers also employ photographs if the ongoing activity is too rapid or there are too many individuals to simultaneously identify. Through repeated surface observations or post-production images, this research reveals features that provide huge increases in discriminability. For example, the ventral fluke patterns of humpback whales (*Megaptera novaeangliae*; Carlson, Mayo, & Whitehead, 1990; Dalla Rosa, Freitas, Secchi, Santos, & Engel, 2004), the lateral body pigmentation of Minke whales (*Balaenoptera acutorostrata*; Arnold, Birtles, Dunstan, Lukoschpk, & Matthews, 2005; Dorsey, Stern, Hoelzel, & Jacobsen, 1990), the scarring of fin whales (*Balaenoptera physalus*; Agler et al., 1990), the natural markings and saddle patches of pilot whales (*Globicephala melas*; Auger-Méthé & Whitehead, 2007), and the callosity patterns of right whales (Würsig & Jefferson, 1990) contain enough variability for accurate individual discrimination from boat based observations. To identify individual bottlenose dolphins, researchers focus on the unique markings and scars present on the dorsal fins and bodies of individuals as well as on the contours of the dorsal fin (Van Hoey, 2013). Maintaining databases of these images allows researchers to reevaluate the features used for identification. For example, in the case of blue whales (*Balaenoptera musculus*) researchers were able to more efficiently identify individuals by including dorsal fin shape in their procedures that previously utilized only body pigmentation (Gendron & Ugalde de la Cruz, 2012).

Identifying individuals. The time-consuming and arduous task of matching photographs of dolphin dorsal fins to known individuals is one of the first tasks that was successfully addressed through automation. Early on, experts systematized their knowledge into a framework for extracting features from

the shape of the dorsal fin for comparison against other known dolphin fins to determine identity (Würsig & Würsig, 1977). This process led to software systems, most notably FinScan (Hillman et al., 2003) and DARWIN (Stanley, 1995; Stewman, Stanley, & Allen, 1995), designed to facilitate this comparison process. This process was still time consuming, as it required that users first trace the contour of the fin, but the work has made automatic contour extraction possible (Hale, 2008; Kreho et al., 1999). Additional work has been done to refine and improve this process, including automatically correcting for orientation and pose (Gilman, Dong, Hupman, Stockin, & Pawley, 2013; Stewman, Debure, Hale, & Russell, 2006).

One of the main difficulties with automating this process is that the background can be noisy and constantly moving. Additionally, depending on the camera or atmospheric conditions, the image may be of poor quality, further reducing the success of identification algorithms. Current methods to identify individual dolphins from their fins are predominantly done in an offline fashion; that is, the images of dolphins during observation are processed after coming back from the field. An opportunity for advancing this technology is to build a real-time system that can process these images, during boat based observations, and label those fins present with the most likely individuals. While these labels in early systems would need to be checked and rechecked in the lab, eventual top-performing systems with minimal error would provide much needed real-time intel for dolphin researchers observing interactive behavior in the wild.

Detection. One difficulty facing such real-time systems lies in the fact that the fins of the individual animals must first be automatically located within an image before a contour can be extracted. Many automated detection methods primarily use motion or color. Because the sea surface is constantly in motion and the colors of the fins and water can often be similar, these methods produce many false detections. The reader is referred to Paiva, Salgado-Kent, Gagnon, Parnum, and McCauley (2015) for active research in this area.

While it is difficult to achieve automated detection of dolphins from surface images or video by visual information alone, it may be easier through infrared video footage, because warm marine mammals stand out from the uniformly cool sea surface and this contrast results in fewer false detections (Graber, 2011). Infrared techniques for marine mammal detection have primarily been used from shore, but have also been used on ships (Zitterbart, Kindermann, Burkhardt, & Boebel, 2013). Work by Graber (2011) demonstrates that infrared video can be used to detect the dorsal fins of mammals from far distances. These methods can also be used to detect blows, as the breath expelled from whales as they surface is warmer than the surrounding water (Graber, 2011; Santhaseelan, Arigela, & Asari, 2012; Santhaseelan & Asari, 2015). Most of this work, however, is focused on larger, migrating whale species.

Aerial Methods

Although having skilled boat based observers is useful for real-time understanding of behavior and much can be learned from these observations at the water's surface, it is difficult to track all animals present. Some animals may not be visible and it can be difficult to determine the activity and arrangement of dolphins. To address these concerns, researchers soon started searching for new vantage points, taking to the skies. In this section, we examine a variety of video collection methods that capture behaviors hidden from surface views and provide avenues to automatically detect and track movement patterns. Although these methods might never provide some of the subtle social behavior present in close-up video footage, they nevertheless can provide researchers with data on behavior and patterns of interaction.

Data collection. Aerial video data collection was originally done via planes (Watkins & Schevill, 1979) and helicopters (Perryman & Lynn, 1994; Scott, Perryman, & Clark, 1985). These approaches, however, were prohibitively expensive and require altitudes that are too high for effective monitoring. Nowacek et al. (2001) introduced an overhead video system consisting of a remotely operated camera attached to the underside of a helium balloon. This system was used for many studies, some of which provided video data on how bottlenose dolphins forage (D. P. Nowacek, 1999), respond to approaching boats (S. M. Nowacek, 1999). Video feeds were immediately stored, monitored for quality and applicability to the study, and adjusted accordingly (Figure 1).

A similar but smaller system, dubbed the "blimp-cam" (Hodgson, 2007), was employed for extensive data collection on the behavior of dugongs (*Dugong dugon*). Lewis, Wartzok, and Heithaus (2011) similarly used a controllable video camera secured beneath a Floatograph© airship, a combination balloon kite system, to investigate leadership within a bottlenose dolphin community. Friedman et al. (2013) improved upon these platforms with the introduction of a low-cost aerial system, consisting of a hybrid parafoil kite and helium balloon that can be elevated to cover the whole study area and required no dedicated operator to orient the camera system.
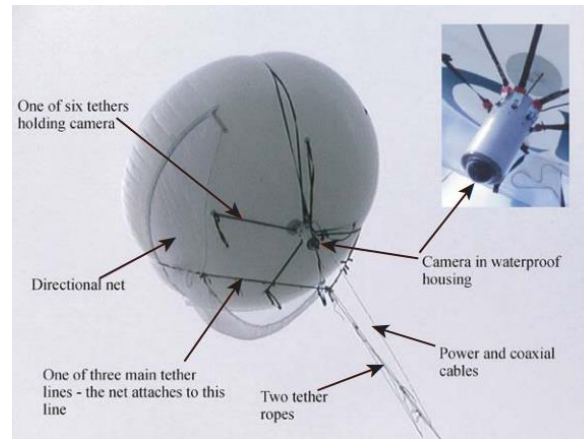

Figure 1. Aerial systems, such as the "blimp-cam" from Hodgson (2007) provide video data on subsurface behavior and group configurations. Smaller balloons combined with kites (Lewis et al., 2011) have replaced older setups that are larger and more expensive (Nowacek et al., 2001).

Drone technology is becoming more mainstream and, as a result, is increasingly being used for the study of marine mammals. The use of unmanned aerial vehicles has several disadvantages however. One difficulty with drones is that automation typically consists of users producing a set of GPS waypoints, which directs the drone to pre-established locations. Whereas this works well for nature conservation surveys over land (van Gemert et al., 2014), it is not ideal over water where the location of subjects is much less predictable. To maintain the proper viewpoint, there must be a dedicated individual, on an accompanying boat, to pilot and maneuver the camera, similar to the manual control present in blimp systems (Nowacek et al., 2001). Some groups have constructed their own proprietary drone systems and developed vision-based control mechanisms, which allow drones to follow an object at a distance given visual estimates (Selby, Corke, & Rus, 2011). However, these control systems are not available for cheap commercial models. The biggest drawback to using drones is that they can stay airborne for only roughly 20-25 min. Until battery technology improves, or unless researchers purchase a collection of drones, researchers will be limited to short recording segments. This does not allow for the recording of long behavioral sequences and limits the flexibility of researchers to deploy systems at the appropriate moments.

Detection. Although these aerial setups provide rich datasets to which continuous and consistent sampling can be applied, there is a lack of methods to deal with the incredible amount of data that is now being generated. Additionally, data collected by aerial techniques is typically processed offline, because algorithms that generate many candidate detections, and use sophisticated deep learning models for classification, require huge amounts of processing power. More lightweight algorithms, such as ones that detect objects by thresholding pixel color values, can run on-board and have shown promise when tracking large single animals against a sea surface that has uniform color (Selby et al., 2011). Other simpler models, which detect objects as collections of parts (deformable parts models) or provide quick classification (support vector machines), still have poor performance when dealing with large groups of

animals (van Gemert et al., 2014). Even in a more successful system that detected half the animals present, roughly 95% of the detections were false positives (Mejias, Duclos, Hodgson, & Maire, 2013). Recent methods using deep learning have improved performance to detect over 80% of animals, but most of the detections (over 70%) still remain errors (Maire, Alvarez, & Hodgson, 2015). Although impressive, this replaces the work of manual detection with the manual elimination of false positives, which may be similarly time consuming.

Individual recognition. Working with aerial photos, recent advances in deep learning have allowed for even more effective automated identification. The National Oceanic and Atmospheric Administration (NOAA), in collaboration with MathWorks and the New England Aquarium, held a competition on Kaggle (Kaggle, 2015) to automate the process of identifying individual right whales (*Eubalaena glacialis*) from aerial photos, as fewer than 500 still remain and most individuals' characteristics are well documented (Bogucki, 2016). The winners of this competition achieved high performance in this task, and because their algorithm is reusable, this will greatly reduce the future manual labor for researchers by automatically identifying right whales in newly acquired photos. This model of providing the questions and data, and then outsourcing the computational work, has become increasingly popular. It leverages the computational resources of many skilled participants and fosters citizen science.

Underwater Methods

Observing marine mammals topside and from above offers a limited viewpoint and it is difficult for researchers to simultaneously track all of the animals in their study and determine the activity and arrangement of animals within a group. Even while tracking a group of dolphins, for example, their fins are not always visible from the surface and overhead cameras cannot always penetrate beneath the surface.

Data collection. Many research efforts routinely survey the ocean (Mellody, 2015; Somerton & Glendhill, 2005), and while these are not necessarily focused on marine mammals, there are exciting opportunities for researchers to use underwater remotely operated vehicles (ROVs) in regions densely populated or frequented by marine mammals. Transects with these devices might help to answer questions about density estimation or activity while targeted observation may be able to contribute meaningfully to studies that require continuous monitoring of ongoing behavior in a specific geographical location. The disadvantages of using ROVs include the underwater noise they create and the possibility that they will disrupt the animals' natural behavior or drive them from the study site.

With the increasing popularity, and decreasing expense, of establishing multi-camera networks, it may become possible to create fixed position systems that track multiple dolphins through an underwater ecosystem, similar to efforts to detect and track fish in coral reefs (Boom et al., 2012). Current research is investigating the advantages, and testing the limits, of such a system that allows for the continuous monitoring of dolphins to study behavior and cognition. For example, the dolphin pools at the Brookfield Zoo were instrumented with 10 cameras directed at underwater portions of the pools and three overhead cameras. These cameras record eight hours of video a day and provide a massive amount of video data that can be used to spearhead research for underwater detection and tracking of marine mammals (Karnowski, Hutchins, & Johnson, 2015).

Studies that use underwater hand-held video recording devices to capture dolphins in their natural habitat and in natural interactions work well as long as the animals are acclimated to humans in the water. For example, over 25 years of video collected while observing dolphins underwater in the Bahamas has led to a broad knowledge base on this species (Cusick & Herzing, 2014; Herzing & Elliser, 2013; Herzing, 2011; Herzing & Johnson, 1997). Recording underwater video also allows for the option of simultaneously recording sound, which is critical for studying communication. However, because such recordings are restricted to the researcher's focus and point of view, they may lack information on other individuals involved in group behavior and make it difficult to track a single focal or group over an extended period.

Automation of detection and localization. Whereas the majority of the underwater detection and tracking efforts have been applied to monitoring fish (e.g., Matai, Kastner, Cutter Jr., & Demer, 2010), the same algorithms represent a great baseline to build upon, provided the cameras are placed farther from the subjects. The first step in these detection algorithms is typically to subtract a static reference image, the background, to find the foreground objects (Figure 2). Background images can be found in various ways, including computing a moving average or using Gaussian Mixture Models (GMMs) (Spampinato, Chen-Burger, Nadarajan, & Fisher, 2008). A recent study on animal detection in camera traps on land utilized Robust Principal Component Analysis (RPCA) (Khorrami, Wang, & Huang, 2012) for background subtraction. RPCA decomposes a video with a static background into a background, a foreground, and an amount of noise. Using GMMs and RPCA for dolphin detection in captivity has allowed researchers to highlight when dolphins move between pools and to locate dolphins within an underwater image (Karnowski et al., 2015). Because underwater scenes containing animals often contain multiple species, especially in the wild, detection systems typically feed into algorithms that classify by species (Cline & Edgington, 2010; Shafait et al., 2016). This may prove useful in dolphin research when recordings include multiple species (Herzing & Johnson, 1997).
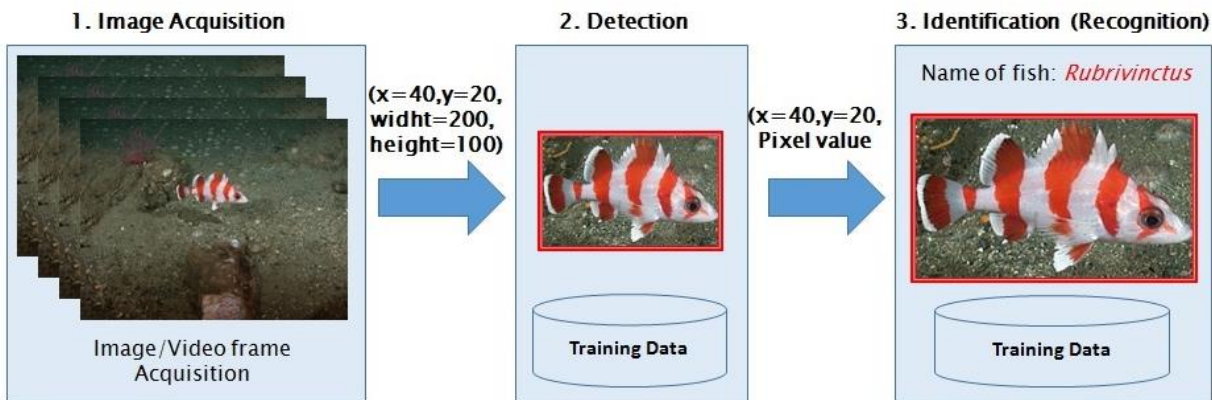

Figure 2. Example system that detects and classifies fish in images (Matai et al., 2010).

Tracking. Underwater tracking algorithms have been centered on the use of the Meanshift algorithm and Camshift algorithm (Spampinato et al., 2008), which use the color histogram of previous frames to predict future target locations. More advanced tracking algorithms, however, take advantage of training data, adapt their object template to changing conditions, and are better at remaining with the current tracked object and not drifting to nearby areas of video. Real-time compressive tracking (CT) (Zhang, Zhang, & Yang, 2012) achieves high success rates on challenging datasets, so it is an ideal candidate for improving the state of the art in underwater tracking, including dolphin tracking. It assumes, however, that the initial tracking window has been specified through human interaction, so to make this fully automated would require algorithms that perform some initial detections.

Emerging Research

There have been many advances in automating the analysis of videos for marine mammal research, and there are many exciting avenues available for future studies. As one of the main ambitions for capturing a wealth of video is to understand marine mammal behavior, methods that automatically uncover activity types and detect patterns within group movements are tantamount to success. Activity recognition, however, is one of the most underdeveloped line of research, as it initially requires solid results in detection and tracking to be successful. Nevertheless, this is an active area of research and there are several labs working on solutions for detecting and classifying behavior in aerial and underwater footage.

Another nascent area of research is individual recognition in underwater video footage. Animals that have prominent markings, such as Atlantic spotted dolphins, can provide salient features from which individuals can be automatically distinguished (T. Starner, personal communication, December 12, 2015). Species with less prominent features, such as bottlenose dolphins, can still be distinguished by trained observers and, thus, the procedure should be amenable to automation. Observational setups, like the one established at the Brookfield Zoo, provide a sufficient number of instances from which algorithms can learn to differentiate individuals. Establishing the proper camera placements and extracting and processing the proper datasets will promote further solutions to this problem.

Some advances to automated video processing may be unexpected, as developments in video capture technologies will likely continue to quickly develop beyond their immediate limits. For example, recent advances in technology have provided prototypes for unmanned vehicles that are equally adept at navigating through the air and underwater (Maia, Soni, & Diez, 2015). With new technology comes the opportunity to re-evaluate, improve, and re-invent the methods built on top of it. There are likely to be new challenges to tackle with automated methods and future algorithmic performance is likely to exceed expectations.

Conclusion

For those who are interested in marine mammal behavior and cognition, video provides a cheap and viable method for systematic and consistent monitoring that has untold benefits as data is analyzed and reanalyzed. Video can be obtained in a variety of ways, including but not limited to, boat based, aerial, and underwater capture methods. With these data, researchers have built automated systems to process and extract information relevant for detecting and tracking freely moving individuals as they engage in social activity. As more sophisticated video capture systems emerge, researchers will continue to tackle issues of how best to mine the wealth of data generated. The future is bright for video surveillance and the development of automated methods to analyze its data, and all those who are interested in furthering the field should seek to tackle open areas of research.

References

Agler, B., Beard, J., Bowman, R., Corbett, H., Frohock, S., … Seipt, I. (1990). Fin whale (*Balaenoptera physalus*) photographic identification: Methodology and preliminary results from the Western North Atlantic. *Report of the International Whaling Commission, 12*, 349–356.

Arnold, P., Birtles, R., Dunstan, A., Lukoschek, V., & Matthews, M. (2005). Colour patterns of the dwarf minke whale *Balaenoptera acutorostrata sensu lato*: Description, cladistic analysis and taxonomic implications. *Memoirs of the Queensland Museum, 51,* 277–307.

Auger-Méthé, M., & Whitehead, H. (2007). The use of natural markings in studies of long-finned pilot whales (*Globicephala melas*). *Marine Mammal Science, 23*, 77–93.

Bogucki, R. (2016). *Which whale is it, anyway? Face recognition for right whales using deep learning.* Retrieved January 16, 2016 from http://deepsense.io/deep-learning-right-whale-recognition-kaggle/

Boom, B., Huang, P., Beyan, C., Spampinato, C., Palazzo, S., …Fisher, R. (2012). Long-term underwater camera surveillance for monitoring and analysis of fish populations. ICPR Workshop on Visual Observation and Analysis of Animal and Insect Behavior (VAIB), Tsukuba Science City, Japan.

Carlson, C., Mayo, C., & Whitehead, H. (1990). Changes in the ventral fluke pattern of the humpback whale (*Megaptera novaeangliae*), and its effect on matching; evaluation of its significance to photo-identification research. *Report of the International Whaling Commission, 12*, 105–111.

Cline, D., & Edgington, D. (2010, August). A detection, tracking, and classification system for underwater images. ICPR Workshop on Visual Observation and Analysis of Animal and Insect Behavior (VAIB), Istanbul, Turkey.

Cusick, J., & Herzing, D. (2014). The dynamic of aggression: How individual and group factors affect the long-term interspecific aggression between two sympatric species of dolphin. *Ethology, 120*, 287–303.

Dalla Rosa, L., Freitas, A., Secchi, E., Santos, M., & Engel, M. (2004, July). *An updated comparison of the humpback whale photo-id catalogues from the Antarctic Peninsula and the Abrolhos Bank, Brazil.* Paper SC/56/SH16 presented to the IWC Scientific Committee, Sorrento, Italy.

Dorsey, E., Stern, S., Hoelzel, A., & Jacobsen, J. (1990). Minke whale (*Balaenoptera acutorostrata*) from the west coast of North America: Individual recognition and small-scale site fidelity. *Report of the International Whaling Commission, 12*, 357–368.

Forster, D. (2002). Consort turnovers as distributed cognition in olive baboons: A systems approach to mind. In M. Bekoff, C. Allen, & G. M. Burghardt (Eds.), *The cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 163–171). Cambridge, MA: MIT Press.

Friedman, W., Potter, S., Krützen, M., Johnson, C., Connor, R., & Hutchins, E. (2013, December). *The aerial advantage: A technical report on the use of three systems for obtaining low-range aerial video of sub-surface behavior among bottlenose dolphins.* Poster session presented at the 20th Biennial Conference of the Biology of Marine Mammals, Dunedin, New Zealand.

Gendron, D., & Ugalde de la Cruz, A. (2012). A new classification method to simplify blue whale photo-identification technique. *Journal of Cetacean Research and Management, 12*, 79–84.

Gilman, A., Dong, T., Hupman, K., Stockin, K., & Pawley, M. (2013). Dolphin fin pose correction using ICP in application to photo-identification. In Rhee, T., Rayudu, R., Hollitt, C., Lewis, J., Zhang, M. (Eds.), *2013 28th International Conference of Image and Vision Computing New Zealand (IVCNZ)* (pp. 388–393). Wellington, New Zealand: IEEE.

Graber, J. (2011). *Land-based infrared imagery for marine mammal detection* (Unpublished doctoral dissertation). University of Washington, Seattle, WA.

Hale, S. (2008). *Unsupervised threshold for automatic extraction of dolphin dorsal fin outlines from digital photographs in DARWIN (Digital Analysis and Recognition of Whale Images on a Network)* (Senior thesis), Eckerd College, St. Petersburg, FL.

Herzing, D. (2011). *Dolphin diaries: My 25 Years with Spotted Dolphins in the Bahamas*. New York, NY: St. Martin's Press.

Herzing, D., & Elliser, C. (2013). Directionality of sexual activities during mixed-species encounters between Atlantic spotted dolphins (*Stenella frontalis*) and bottlenose dolphins (*Tursiops truncatus*). *International Journal of Comparative Psychology, 26,* 124–134.

Herzing, D., & Johnson, C. (1997). Interspecific interactions between Atlantic spotted dolphins (*Stenella frontalis*) and bottlenose dolphins (*Tursiops truncatus*) in the Bahamas, 1985–1995. *Aquatic Mammals, 23*, 85–99.

Herzing, D., & Johnson, C. (2015). Dolphin communication and cognition: Past, present, and future. Cambridge, MA: MIT Press.

Hillman, G., Würsig, B., Gailey, G., Kehtarnavaz, N., Drobyshevsky, A., …Weller, D. (2003). Computer-assisted photo-identification of individual marine vertebrates: A multi-species system. *Aquatic Mammals, 29*, 117–123.

Hodgson, A. (2007). "BLIMP-CAM": Aerial video observations of marine animals. *Marine Technology Society Journal, 41*, 39–43.

Johnson, C. (2001). Distributed primate cognition: A review. *Animal Cognition, 3*, 167–183.

Johnson, C. (2010). Observing cognitive complexity in primates and cetaceans. *International Journal of Comparative Psychology, 23,* 587–624.

Johnson, M., & Tyack, P. (2003). A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE Journal of Oceanic Engineering, 28*, 3–12.

Karnowski, J., Hutchins, E., & Johnson, C. (2015, January). Dolphin detection and tracking. Paper presented at the Workshop on Automated Analysis of Video Data for Wildlife Surveillance, Waikoloa Beach, HI.

Khorrami, P., Wang, J., & Huang, T. (2012). Multiple animal species detection using robust principal component analysis and large displacement optical flow. ICPR Workshop on Visual Observation and Analysis of Animal and Insect Behavior (VAIB). Tsukuba Science City, Japan.

Kreho, A., Kehtarnavaz, N., Araabi, B., Hillman, G., Würsig, B., & Weller, D. (1999). Assisting manual dolphin identification by computer extraction of dorsal ratio. *Annals of Biomedical Engineering, 27*, 830–838.

Lewis, J. S., Wartzok, D., & Heithaus, M. (2011). Highly dynamic fission–fusion species can exhibit leadership when traveling. *Behavioral Ecology and Sociobiology, 65*, 1061–1069.

Maia, M., Soni, P., & Diez, F. (2015). Demonstration of an aerial and submersible vehicle capable of flight and underwater navigation with seamless air-water transition. Retrieved from https://arxiv.org/abs/1507.01932

Maire, F., Alvarez, L., & Hodgson, A. (2015). Automating marine mammal detection in aerial images captured

during wildlife surveys: A deep learning approach. In 28th Australasian Joint Conference on Artificial Intelligence (pp. 379–385), Springer International Publishing

Matai, J., Kastner, R., Cutter Jr., G., & Demer, D. (2010). *Automated techniques for detection and recognition of fishes using computer vision algorithms*. Report of the National Marine Fisheries Service Automated Image Processing Workshop (NOAA Technical Memorandum NMFS-F/SPO-121). Retrieved from https://swfsc.noaa.gov/publications/CR/2012/2012Matai.pdf

McCowan, B., Hanser, S., & Doyle, L. (1999). Quantitative tools for comparing animal communication systems: Information theory applied to bottlenose dolphin whistle repertoires. *Animal Behaviour, 57*, 409–419.

Mejias, L., Duclos, G., Hodgson, A., & Maire, F. (2013, September). *Automated marine mammal detection from aerial imagery.* Paper presented at 2013 OCEANS-San Diego, San Diego, CA. Retrieved from http://ieeexplore.ieee.org/abstract/document/6741088/

Mellody, M. (2015). *Robust methods for the analysis of images and videos for fisheries stock assessment: Summary of a workshop*. Washington, DC: National Academies Press.

Nowacek, D. P. (1999). *Sound use, sequential behavior and ecology of foraging bottlenose dolphins, Tursiops truncatus* (Unpublished doctoral dissertation). Massachusetts Institute of Technology and Woods Hole Oceanographic Institution, Boston, MA.

Nowacek, D. P., Tyack, P., & Wells, R. (2001). A platform for continuous behavioral and acoustic observation of free-ranging marine mammals: Overhead video combined with underwater audio. *Marine Mammal Science, 17*, 191–199.

Nowacek, S. M. (1999). *The effects of boat traffic on bottlenose dolphins, Tursiops truncatus, in Sarasota Bay, Florida* (Unpublished master's thesis). University of California Santa Cruz, Santa Cruz, CA.

Paiva, E., Salgado-Kent, C., Gagnon, M., Parnum, I., & McCauley, R. (2015). An assessment of the effectiveness of high definition cameras as remote monitoring tools for dolphin ecology studies. *PloS one, 10*, e0126165.

Perryman, W., & Lynn, M. (1994). Examination of stock and school structure of striped dolphin (*Stenella coeruleoalba*) in the eastern Pacific from aerial photogrammetry. *Fishery Bulletin, 92*, 122–131.

Right whale recognition. (2015, August 27). In *Kaggle*. Retrieved from https://www.kaggle.com/c/noaa-right-whale-recognition

Santhaseelan, V., Arigela, S., & Asari, V. (2012). Neural network based methodology for automatic detection of whale blows in infrared video. In George Bebis, G., Boyle, R., Parvin, B., Koracin, D., Fowlkes, C., Wang, S., Choi, M., Mantler, S., Schulze, J., Acevedo, D., Mueller, K., Papka, M. (Eds.), *Advances in Visual Computing* (pp. 230–240). Berlin Heidelberg: Springer.

Santhaseelan, V., & Asari, V. (2015). Automated whale blow detection in infrared video. In Zhou, J., Bai, X., Caelli, T. (Eds.), *Computer Vision and Pattern Recognition in Environmental Informatics* (pp. 58–78). Hershey, PA: Information Science Reference.

Scott, M., Perryman, W., & Clark, W. (1985). The use of aerial photographs for estimating school sizes of cetaceans. *Inter-American Tropical Tuna Commission Bulletin, 18*, 381–419.

Selby, W., Corke, P., & Rus, D. (2011). *Autonomous aerial navigation and tracking of marine animals.* Paper presented at the Australian Conference on Robotics and Automation (ACRA), Melbourne, Australia. Retrieved from http://www.araa.asn.au/acra/acra2011/papers/pap165.pdf

Shafait, F., Mian, A., Shortis, M., Ghanem, B., Culverhouse, P., …Harvey, E. (2016). Fish identification from videos captured in uncontrolled underwater environments. *ICES Journal of Marine Science.* Advance online publication. Retrieved from http://icesjms.oxfordjournals.org/content/early/2016/07/14/icesjms.fsw106.abstract

Somerton, D., & Glendhill, C. (2005). Report of the National Marine Fisheries Service Workshop on Underwater Video Analysis. (NOAA Technical Memorandum NMFS-F/SPO-68). Retrieved from http://spo.nmfs.noaa.gov/tm/tm68.pdf

Spampinato, C., Chen-Burger, Y., Nadarajan, G., & Fisher, R. (2008). Detecting, tracking and counting fish in low quality unconstrained underwater videos. In Ranchordas, A., Araújo (Eds.), *VISAPP 2008: Proceedings of the Third International Conference on Computer Vision Theory and Applications* (pp. 514–519). Funchal, Portugal: INSTICC

Stanley, R. (1995). Darwin: Identifying dolphins from dorsal fin images (senior thesis). Eckerd College, St. Petersburg, FL.

Stewman Stewman, J., Debure, K., Hale, S., & Russell, A. (2006). Iterative 3-D pose correction and content-based image retrieval for dorsal fin recognition. In Campilho, A., Kamel, M. (Eds.), *Image analysis and recognition* (pp. 648–660). Berlin, Germany: Springer.

Stewman, J., Stanley, R., & Allen, M. (1995). DARWIN: A system to identify dolphins from fin profiles in digital

images. *8th Annual Florida Artificial Intelligence Research Symposium (FLAIRS)*. Melbourne, FL: AAAI.

Tyson, R., Friedlaender, A., Ware, C., Stimpert, A., & Nowacek, D. P. (2012). Synchronous mother and calf foraging behaviour in humpback whales *Megaptera novaeangliae*: Insights from multi-sensor suction cup tags. *Marine Ecology Progress Series*, *457*, 209–220.

van der Hoop, J., Fahlman, A., Hurst, T., Rocho-Levine, J., Shorter, K., …Moore, M. J. (2014). Bottlenose dolphins modify behavior to reduce metabolic effect of tag attachment. *Journal of Experimental Biology, 217*, 4229–4236.

van Gemert, J., Verschoor, C., Mettes, P., Epema, K., Koh, L., & Wich, S. (2014). Nature conservation drones for automatic localization and counting of animals. In Agapito, L., Bronstein, M., Rother, C. (Eds.), In Computer Vision-ECCV 2014 Workshops (pp. 255–270). Berlin, Germany: Springer International Publishing.

Van Hoey, N. (2013). *Photo-identification and distribution of bottlenose dolphins (Tursiops truncatus) off Bimini, The Bahamas, 2006-2009* (Unpublished master's thesis). Alaska Pacific University, Anchorage, AK.

Watkins, W., & Schevill, W. (1979). Aerial observation of feeding behavior in four baleen whales: *Eubalaena glacialis, Balaenoptera borealis, Megaptera novaeangliae,* and *Balaenoptera physalus. Journal of Mammalogy, 60*, 155–163.

Würsig, B., & Jefferson, T. (1990). Methods of photo-identification for small cetaceans. In Hammond, P., Mizroch, S., Donovan, G. (Eds.), *Individual recognition of cetaceans: Use of photo-identification and other techniques to estimate population parameters* (pp. 43–51). Cambridge, MA: The International Whaling Commission.

Würsig, B., & Würsig, M. (1977). The photographic determination of group size, composition, and stability of coastal porpoises (*Tursiops truncatus*). *Science, 198*, 755–756.

Zhang, K., Zhang, L., & Yang, M. (2012, October). Real-time compressive tracking. In Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (Eds.). Paper presented at European Conference on Computer Vision (pp. 864-877). Berlin, Germany: Springer.

Zitterbart, D., Kindermann, L., Burkhardt, E., & Boebel, O. (2013). Automatic round-the-clock detection of whales for mitigation from underwater noise impacts. *PloS one, 8*, e71217.