2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Detection of Dugongs from Unmanned Aerial Vehicles

Frederic Maire[a], Luis Mejias[a,b], Amanda Hodgson[c], Gwenael Duclos[d]

Abstract— **Monitoring and estimation of marine populations is of paramount importance for the conservation and management of sea species. Regular surveys are used to this purpose followed often by a manual counting process. This paper proposes an algorithm for automatic detection of dugongs from imagery taken in aerial surveys. Our algorithm exploits the fact that dugongs are rare in most images, therefore we determine regions of interest partially based on color rarity. This simple observation makes the system robust to changes in illumination. We also show that by applying the extended-maxima transform on red-ratio images, submerged dugongs with very fuzzy edges can be detected. Performance figures obtained here are promising in terms of degree of confidence in the detection of marine species, but more importantly our approach represents a significant step in automating this type of surveys.**

I. Introduction

The conservation and management of many marine mammal (whale, dolphin and dugong) populations relies on accurate and precise estimates of their abundance, distribution and habitat use. The estimation of a population and its geographical distribution is not only important for gaining understanding of particular species, but often government regulations impose strict requirements to industry working in the vicinity of their habitat.

In Australia, regular surveys have been conducted since the 80s, most notably in Queensland and Torres Strait [1][2], and since the 1990s in Shark Bay and Exmouth [3][4][5][6]. Whales [7] and Sea Lions [8] have also been monitored using aerials surveys. In the US, the Marine Mammal Protection Act (MMPA) of 1972, requires an annual stock assessment of all marine mammal species in US waters. Many of these stock assessments and the consequential management actions to conserve marine mammals are based on minimum population estimates from aerial surveys. In Europe [9] and Canada [10] abundance estimates of cetaceans also rely on aerial surveys. The datasets produced from aerial surveys form the basis of many studies to determine the ecological requirements of species [11][12][13][14], they are also used to assess the effectiveness of marine mammal sanctuaries [15].

There exist already procedures and standards in place for conducting marine mammal aerial surveys. They often require manned aircraft with specialised equipment and onboard crew for manual counting of the species. This process is time consuming and requires very specialised skills for the identification of species in the data collected. Therefore, automation of whole or part of this process would greatly benefit researchers and hence the conservation of marine mammal species.

In this paper we propose an approach to automatically detect marine mammals in aerial imagery taken using custom payloads onboard aircraft, and hence contributing to overcome the limitations of current data analysis. We derive a novel pattern recognition algorithm that exploits specific features of the environment.

This paper is structured as follows. Section II reviews recent work. Section III describes the approach developed. Section IV-A outlines the data used in the experiments. Section IV-B presents the outcomes and analysis of data. Finally, section V describes some of the lessons learnt and future work planned.

II. Background

The detection and monitoring of marine mammal species can be conducted using a variety of different sensors complementary in the type

[a] School of Electrical Engineering and Computer Science, Queensland University of Technology, Brisbane, Australia. f.maire@qut.edu.au

[b] Australian Research Centre for Aerospace Automation Brisbane, Australia. luis.mejias@qut.edu.au

[c] Murdoch University Cetacean Research Unit, Murdoch University. A.Hodgson@murdoch.edu.au

[d] Wildlife Image Processing Solutions for Environmental Assessments, gwenael.duclos@gmail.com

of information they provide. For instance, active sensors (such as radar or sonar) are used to detect marine species [16][17]. They offer robustness against environmental conditions, however they tend to be complex and computationally intensive in terms of signal processing.

Passive sensors (such as imagery or acoustic) are also used in surveys. Acoustic sensors rely on mammal vocalisations and a low ambient noise so the sound is distinguishable from the background. Acoustic approaches are often less than desirable because they introduce anthropomorphic noise into the marine environment potentially affecting cetacean behavior [18]. Imagery or visual observation approaches are limited by atmospheric conditions and greatly affected by illumination conditions [19], [20], [21]. Infrared (IR) imagery has the potential to improve upon visual observations by enabling nighttime detections [22], [23]. Despite its limitations, visual imagery or vision is an attractive solution given it offers a rich source of information. Cameras are inexpensive and low power consumption which provides great advantages considering the inherent limitations in size, weight and power of many light aircraft or small/medium Unmanned Aerial Vehicles (UAVs).

III. Detection approach

What makes the detection of dugongs particularly challenging is that their appearance varies dramatically with the sea conditions. Their apparent color changes with the depth and the turbidity of the water. Although the shape of a dugong is relatively rigid, their tail is not always visible. Moreover, parts of their bodies can be covered by small waves with breaking crests or whitecaps.

The strategy that we adopted consists in two main stages. In the first stage, we determine regions of interest through a number of color and morphological filters (details in Section III-A). In the second stage, the regions of interest produce a number of candidate blobs obtained by local segmentation (details in Section III-B). The shape of these blobs is then analysed with geometric features and later compared with a small set of shape templates.

A. Regions of interest determination

In a nutshell, a blob is of interest if it is salient, has an interesting color and is not created by whitecaps. Experimentally, we discovered that the blobs corresponding to dugongs are locally maximal plateau regions in the scalar image I_{rr} derived from the original color image by computing for each pixel (R, G, B) the ratio $\rho = \frac{R}{G+B}$. We call ρ the *red-ratio* of the pixel.

1) Red-ratio computation: The bottom right plot of Figure 1 is the red-ratio image of the bottom left image. The bottom left image was derived from the original image (top left) by identifying the whitecaps (top right) and inpainting them (bottom left). That is, for each color channel, the whitecaps are replaced by smoothly interpolating inward from the pixel values on the boundary of the whitecaps by solving Laplace's equation. In Matlab®, this can be done by calling the *roifill* function on each color channel. The benefit of inpainting the whitecaps is clearly visible when comparing Figure 3, the surface plot of the red-ratio of the image in Figure 2, and Figure 4, the surface plot of the red-ratio of the same image but with the whitecaps inpainted. The red top of the central plateau in Figure 4 is smoother and flatter than the one in Figure 3. The steps leading to the computation of the red-ratio image I_{rr} are spread from Line 2 to Line 5 in Algorithm 1.

A pixel $p = (v_r, v_g, v_b)$ is considered to be part of a whitecaps if for each color channel $c \in \{R, G, B\}$, the pixel value of channel c satisfies $v_c > 1.3 \times \mu_c$ where μ_c denotes the mean of the color channel c in the image. The factor 1.3 was chosen empirically and its exact value is not a critical factor. The binary image I_{wc} is dilated into image I_{wcd} with a disk of radius 5 as the structuring element.

2) Pixel rarity and entropy filtering: Another strong clue that a dugong is present is the rarity of the color of a blob. To approximate the (R, G, B) probability distribution of the pixels in a given image, we compute a $20 \times 20 \times 20$ frequency table of the (R, G, B) triplets occuring in the image. A triplet (R, G, B) is considered *rare*, if the probability mass of the associated cell in the frequency table is less than 0.02. Figure 8 shows that the 3 dugongs in the image of Figure 5 have

a rare color.

There are many locally maximal plateaus in the red-ratio image that occur by chance in the sea background. To eliminate most of them, we compute a *seed image* I_s^2 (Line 8 to Line 10 of Algorithm 1). Plateaus in the red-ratio image that do no meet any seed pixels are discarded. In Figure 8, the dugongs appear clearly as local maxima of the red-ratio image.

The binary seed image I_s^2 captures the regions of pixels with rare color that also contains some high entropy pixels. Here, the *pixel entropy* denotes the entropy of its 9×9 neighborhood. The entropy of a faint edge pixel of a dugong below the surface is typically above 4.2. Figures 5 and 6 show that the seed regions cover the dugongs.

To determine the plateaus of the red-ratio image, we compute the *extended-maxima transform* [24] with threshold t taking 8 values logarithmically spaced between 0.02 and 0.2 (Lines 11 and 12 of Algorithm 1).

3) Morphological computation of the plateaus of I_{rr}: As can be seen from Figure 4, the plateaus of I_{rr} corresponding to dugongs can have a blunted and shallow relief for underwater dugongs (not much difference with the surroundings) or can be very sharp and high for dugongs on the surface of the water (strong contrast with the surroundings). These subtle variations in red-ratio can happen at any level. It is theoretically possible to thresdhold I_{rr} with many values, but this would be computationally very expensive. A more efficient approach is to apply the extended-maxima transform.

The extended-maxima builds on the *t-maxima transform* ([24, pp. 170-171]). Conceptually, the t-maxima transform removes all local maxima whose height with respect to their (lower) neighbors is less than t. In other words, every peak that stands out by less than t disappears as if pushed down to the level of its neighbors. Then, the regional maxima of the t-maxima transform are computed. Regional maxima are connected components of pixels with a constant intensity value, and whose external boundary pixels all have a lower value. The t-plateaus on Line 12 in Algorithm 1 are the regional maxima of the t-maxima transform.

input : I_{rgb} a RGB color image
output: L a list of blob bounding boxes believed to contain dugongs

1 **begin**
2 $I_{wc} \longleftarrow$ whitecaps image (binary)
3 $I_{wcd} \longleftarrow$ dilatation of I_{wc}
4 $I_{ip} \longleftarrow$ inpaint I_{rgb} with mask I_{wcd}
5 $I_{rr} \longleftarrow$ red ratio of I_{ip}
6 $I_f \longleftarrow$ color frequency image
7 $I_{he} \longleftarrow$ high (> 4.2) entropy regions (binary)
8 $I_s^0 \longleftarrow$ rare color pixels ($I_f < 0.02$) not in I_{wcd}
9 $I_s^1 \longleftarrow$ blobs of I_s^0 intersecting I_{he}
10 $I_s^2 \longleftarrow$ eliminate small blobs from I_s^1
11 **for** $t \in \text{logspace}(0.02, 0.2, 8)$ **do**
12 $I_t^0 \longleftarrow$ t-plateau regions of I_{rr}
13 $I_t^1 \longleftarrow$ blobs of I_t^0 intersecting I_s^2
14 $I_t^2 \longleftarrow$ blobs of I_t^1 of right size
15 $I_t^3 \longleftarrow$ blobs of I_t^2 with low proportion of I_{wcd}
16 **for** *blob in* I_t^3 **do**
17 **if** *hasDugongShape(blob)* **then**
18 Add blob to L
19 **end**
20 **end**
21 **end**
22 **end**

Algorithm 1: Dugong Detector Algorithm

B. Shape analysis

When scanning the blobs of I_t^3 (Line 15 of Algorithm 1), we perform a shape analysis on the blob itself and on a *twin* blob obtained by local binary segmentation of the 30×30 window centered at the centroid of the first blob.

1) Local segmentation: The local segmentation of the 30×30 window is performed on the grey image in the four quadrants using Otsu method [25]. The four segmented binary images are then merged after flipping labels if necessary to ensure that the core blob receives consistent labels. This situation can arise when one part of the dugong is on a light background (sand for example), and the rest of the dugong is on a darker background (seagrass for example).

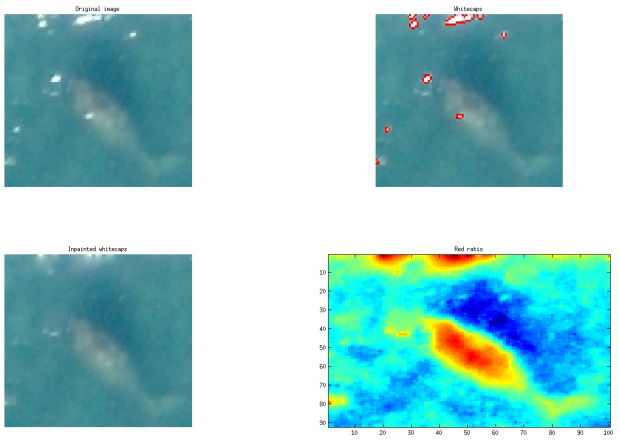

Fig. 1. Top left: original window. Top right: whitecaps. Bottom left: inpainted whitecaps. Bottom right: red-ratio of bottom left image.

Fig. 2. A dugong partially covered with whitecaps. The dugong's tail is on the right.

2) Shape features: The shapes of a blob and its twin are tested in the same way. A number of geometric measurements are extracted into a feature vector. Then this vector is classified using a hand coded decision list (comparing combination of measurements to thresholds). A blob is more likely to relate to a dugong if its shape is elliptical. A good measure of this property is the following ratio

$$\frac{\pi \times \text{MajorAxisLength} \times \text{MinorAxisLength}}{4 \times \text{Area}}$$

The closer to 1 is this ratio, the more elliptical the shape of the blob is. We call this ratio the *elliptic ratio*. The feature vector used for shape classification includes a template similarity measure, the blob diameter, the length of major and minor axes of the blob, and the elliptic ratio.

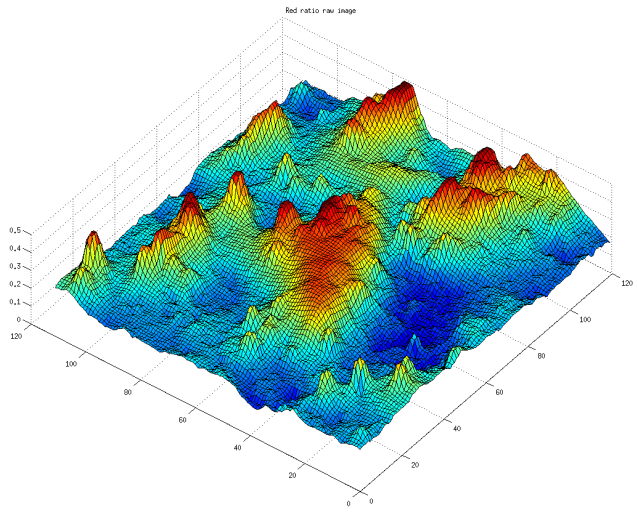

Fig. 3. Red-ratio image of the raw image of Figure 2

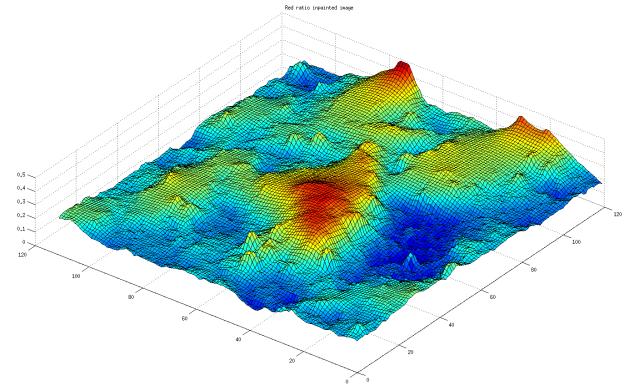

Fig. 4. Red-ratio image of the inpainted image of Figure 2

IV. Experimentation and Analysis

A. Data collection and experiment setup

Our testing dataset consisted of pictures captured in Shark Bay (Western Australia) using a UAV during seven flights. Onboard the UAV a Nikon 12 megapixel digital SLR camera was mounted downward-looking with a standard 50 mm lens and a polarising filter. Each image was tagged in real-time with GPS information from a dedicated receiver. During each flight a set of 10 transects were flown at three different altitudes: 500 ft., 750 ft. and 1000 ft. Transects were designed to cover different habitats (i.e. open

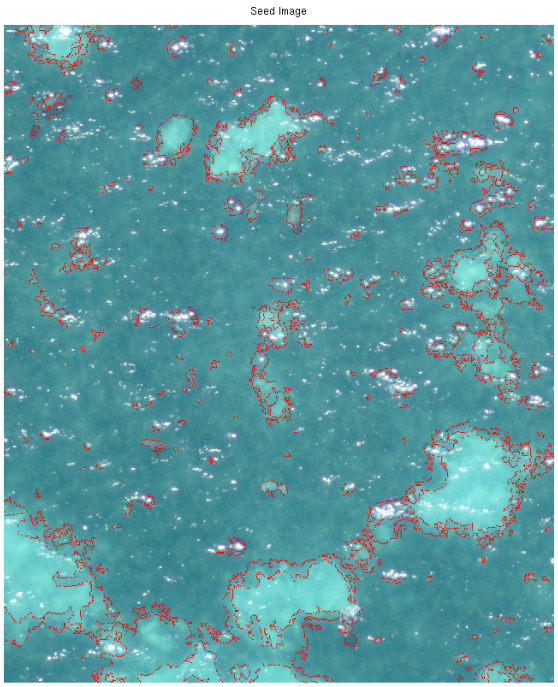

Fig. 5. Contours of the seed regions

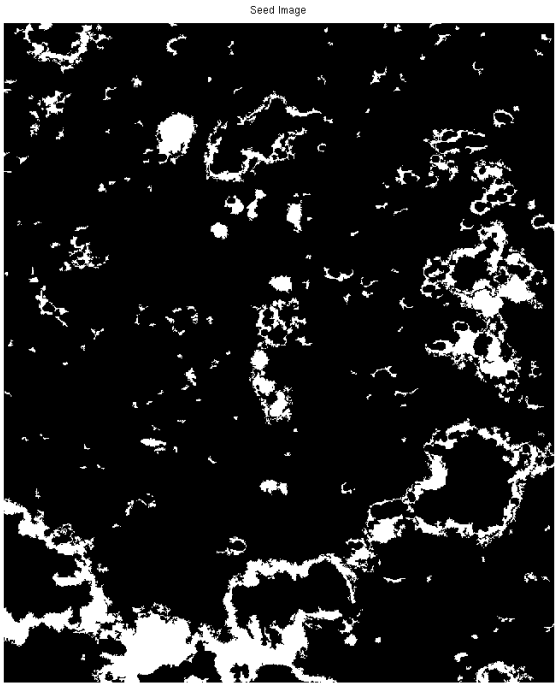

Fig. 6. Binary represention of the seed regions (an alternative view of Figure 5)

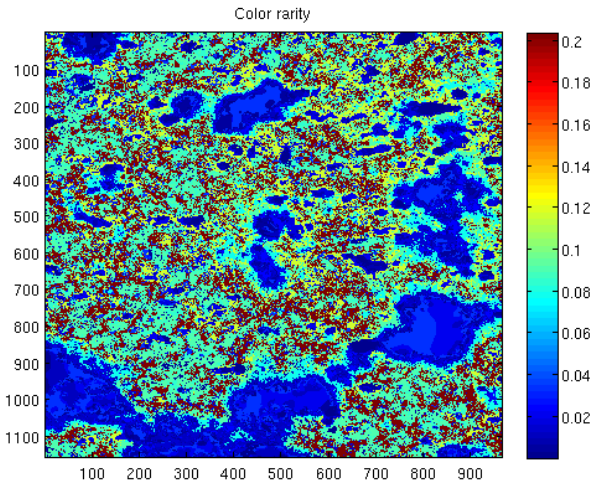

Fig. 7. Color rarity (from the same image as Figure 5)

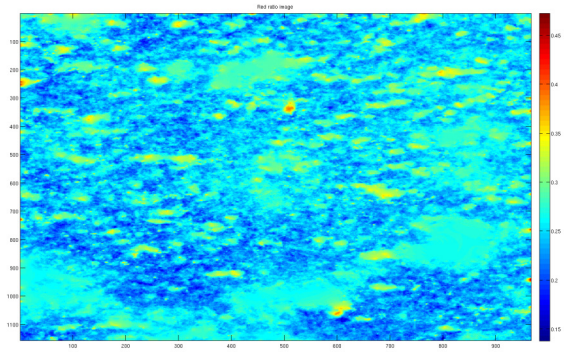

Fig. 8. Red-ratio (from the same image as Figure 5)

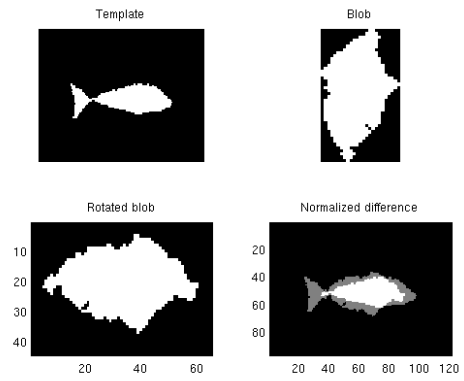

Fig. 9. Before comparing a blob (top right) to a template (top left), the blob is normalized by rotating it so that its principal axis is horizontal (bottom left), and rescaling it so that its area is the same as the template. Then the ratio of the intersection set over the union set of the two shapes produces a similarity score.

water and sea grass banks) in an area where large numbers of dugongs were expected to occur, and were performed in different sea state conditions. Sea condition were defined using the Beaufort scale [26].

An evaluation dataset with 28 pictures of resolution 4288 by 2848 (equivalent to 1113 VGA pictures) was used to represent the variability of all environmental conditions. These 28 pictures were manually labelled by a marine biologist.

B. Analysis

The lack of published benchmarks for algorithm of this type only allows to draw conclusions in a limited context. A preliminary and unpublished version of this approach was used to evaluate the performance of the current algorithm. Previous versions were based on color thresholding in the HSV space for segmentation and blob profile measurements for shape classification. The performance of this base-line system was as follows; The recall was 51.4% and the precision was 4.97%. The large number of false positives explains the low score for precision. The system presented in this paper was tested on the same dataset. These previous results are improved with a recall of 69.4% and a precision of 30%. As expected, the performance of the system is very sensitive to the sea conditions. In calm sea conditions, like in Figure 10, the system performs very well. But as the sea surface becomes rougher the performance of the system degrades (Figures 11 and 12). The performance of the system restricted to calm conditions (one third of the dataset) is significantly better with a recall of 75.4% and a precision of 87.5%, which is comparable to non-expert human performance. While it is difficult to benchmark to the human visual in identifying same type of marine mammal in images, performance figures obtained by our approach are promising in terms of degree of confidence in detection dugongs from aerial images. Further efforts will involve creating a baseline benchmark for quantitative evaluation.

V. Conclusion

This paper introduces a number of features (like the rarity of a color and the red-ratio) to help determine regions of interest. The application of the extended-maxima transform allows us to detect dugongs hardly visible to the naked eye. When the sea condition is mild, the performance of the presented system is satisfactory for practical purpose. Robustness with respect to illumination is achieved by the combination of color rarity and extended-maxima transform. However, the system still requires some improvement (mainly to reduce the number of false positives) when there are breaking waves.

With our detection system, all dugongs present

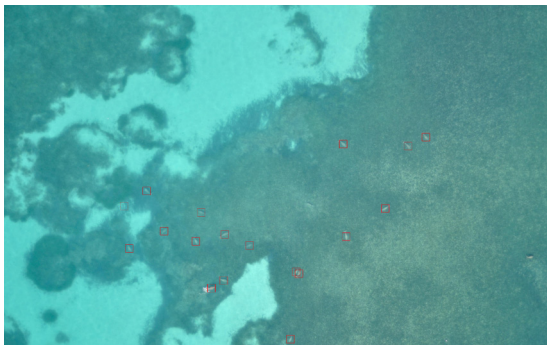

Fig. 10. Result image 8739: 17 out 19 dugongs detected, 1 false positive.

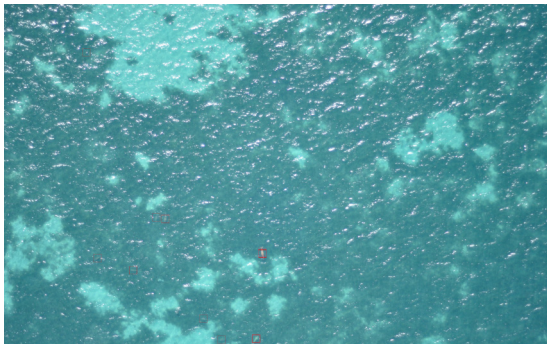

Fig. 11. Result image 0774: 7 out 13 dugongs detected, no false positives.

Fig. 12. Result image 0241: 3 out 4 dugongs detected, 14 false positives.

in an image are tagged as regions of interest. Unfortunately, the shape filtering module misclassifies some of them. We hope to further improve this module by replacing our hand-coded shape classifier with a learnt one (a neural network or a support vector machine) fed with the same feature vector as the one described in Section III-B.2.

REFERENCES

[1] H. Marsh, I. R. Lawler, D. Kwan, S. Delean, K. Pollock, and M. Alldredge, "Aerial surveys and the potential biological removal techniques indicate that the Torres Strait dugong fishery is unsustainable," *Animal Conservation*, vol. 7, pp. 435–443, 2004.

[2] H. Marsh and I. R. Lawler, "Dugong distribution and abundance on the urban coast of Queensland: a basis for management. pp 79," Marine and Tropical Science Research Facility Interim Projects 2005-06 Final Report, Project 2,, Townsville, Queensland, Tech. Rep., 2006.

[3] H. Marsh, R. I. T. Prince, W. K. Saalfeld, and R. Shepherd, "The distribution and abundance of the dugong in Shark Bay, Western Australia," *Wildlife Research*, vol. 21, pp. 149–161, 1994.

[4] A. R. Preen, H. Marsh, I. R. Lawler, R. I. T. Prince, and R. Shepherd, "Distribution and abundance of dugongs, turtles, dolphins and other megafauna in Shark Bay, Ningaloo Reef and Exmouth Gulf, Western Australia," *Wildlife Research*, vol. 24, pp. 185–208, 1997.

[5] N. J. Gales, R. D. McCauley, J. M. Lanyon, and D. K. Holley, "Change in abundance of dugongs in Shark Bay, Ningaloo and Exmouth Gulf, Western Australia: evidence for large scale migration," *Wildlife Research*, vol. 31, pp. 283–290, 1994.

[6] D. K. Holley, I. R. Lawler, and N. J. Gales, "Summer survey of dugong distribution and abundance in Shark Bay reveals additional key habitat area," *Wildlife Research*, vol. 33, no. 3, pp. 243–250, 2006.

[7] M. J. Noad, R. A. Dunlop, D. Paton, and D. H. Cato, "An update of the east Australian humpback whale population (e1) rate of increase." International Whaling Commission Scientific Committee, Tech. Rep., 2008.

[8] P. D. Shaughnessy, T. E. Dennis, and P. G. Seager, "Status of Australian sea lions, Neophoca cinerea, and New Zealand fur seals, Arctocephalus forsteri, on Eyre Peninsula and the far west coast of South Australia." *Wildlife Research*, vol. 32, no. 1, pp. 85–181, 2005.

[9] P. S. Hammond, P. Berggren, H. Benke, D. L. Borchers, A. Collet, M. P. Heide-Jorgensen, S. Heimlich, A. R. Hiby, M. F. Leopold, and N. Oien, "Abundance of harbour porpoise and other cetaceans in the north sea and adjacent waters," *Journal of Applied Ecology*, vol. 39, no. 2, pp. 361–376, 2002.

[10] M. C. S. Kingsley and R. R. Reeves, "Aerial surveys of cetaceans in the Gulf of St. Lawrence in 1995 and 1996," *Canadian Journal of Zoology*, vol. 76, no. 8, pp. 426–445, 1996.

[11] B. A. Craig and J. E. Reynolds, "Determination of manatee population trends along the Atlantic coast of Florida using a Bayesian approach with temperature-adjusted aerial survey data," *Marine Mammal Science*, vol. 20, no. 3, pp. 386–400, 2004.

[12] C. A. Keller, L. I. Ward-Geiger, W. B. Brooks, C. K. Slay, C. R. Taylor, and B. J. Zoodsma, "North Atlantic right whale distribution in relation to sea-surface temperature in the southeastern United States calving grounds," *Marine Mammal Science*, vol. 22, no. 2, pp. 426–445, 2006.

[13] R. P. Sonntag, A. R. H. H. Benke, R. Lick, and D. Adelung, "Identification of the first harbour porpoise (phocoena phocoena) calving ground in the north sea." *Journal of Sea Research*, vol. 41, no. 3, p. 225, 232 1999.

[14] M. F. Baumgartner, "The distribution of Risso's dolphin (Grampus griseus) with respect to the physiography of the northern Gulf of Mexico," *Marine Mammal Science*, vol. 13, no. 4, pp. 614–638, 1997.

[15] E. Slooten, W. Rayment, and S. Dawson, "Offshore distribution of Hector's dolphins at Banks Peninsula, New Zealand: is the Banks Peninsula Marine Mammal sanctuary large enough?" *New Zealand Journal of Marine Fresh Water*, vol. 40, no. 2, pp. 333–343, 2006.

[16] D. DeProspo, J. Mobley, W. Hom, and M. Carron, "Radar-based detection, tracking and speciation of marine mammals from ships," Arete Associates Project Report, Tech. Rep., 2004.

[17] S. Anderson and J. Morris, "On the detection of marine mammals with ship-borne polarimetric microwave radar," in *OCEANS 2010 IEEE - Sydney*, May, pp. 1–6.

[18] D. McGaughey, D. Marcotte, M. Korenberg, and J. Theriault, "Detection and classification of marine mammal clicks," in *OCEANS 2010*, Sept., pp. 1–10.

[19] Y. Podobna, J. Sofianos, J. Schoonmaker, D. Medeiros, C. Boucher, D. Oakley, and S. Saggese, "Airborne multispectral detecting system for marine mammals survey," in *Proc. SPIE 7678, Ocean Sensing and Monitoring II, 76780G.*, April 20, 2010 2010. [Online]. Available: http://dx.doi.org/10.1117/12.849485

[20] Y. Podobna, J. Schoonmaker, C. Boucher, and D. Oakley, "Optical detection of marine mammals," in *Proc. SPIE 7317, Ocean Sensing and Monitoring*, vol. 7317, 2009.

[21] W. Selby, P. Corke, and D. Rus, "Autonomous aerial navigation and tracking of marine animals," in *Proc. of the Australian Conference on Robotics and Automation (ACRA)*, 2011.

[22] J. Schoonmaker, T. Wells, G. Gilbert, Y. Podobna, I. Petrosyuk, and J. Dirbas, "Spectral detection and monitoring of marine mammals," in *SPIE 6946, Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications V, 694606*, 2008. [Online]. Available: http://dx.doi.org/10.1117/12.777740

[23] J. Schoonmaker, Y. Podobna, C. Boucher, J. Sofianos, D. Oakley, D. Medeiros, and J. Lopez, "The utility of automated electro-optical systems for measuring marine mammal densities," in *OCEANS 2010*, Sept., pp. 1–6.

[24] P. Soille, *Morphological Image Analysis: Principles and Applications*. Springer-Verlag, 1999.

[25] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, pp. 285–296, 1975.

[26] S. Huler, *Defining the Wind: The Beaufort Scale and How a 19th-Century Admiral Turned Science into Poetry*. Crown, 2004.