Automated Marine Mammal Detection From Aerial Imagery

Luis Mejias[a], Gwenael Duclos[b]
[a]Australian Research Centre for Aerospace Automation
Queensland University of Technology
Brisbane, Australia
[b]Wildlife Image Processing
Solutions for Environmental Assessments
Email: luis.mejias@qut.edu.au, gwenael.duclos@gmail.com

Amanda Hodgson[c], Frederic Maire[d]
[c] Murdoch Cetacean Research Unit
Murdoch University
Murdoch, Australia
[d]CyPhy Lab
School of Electrical Engineering and Computer Science
Queensland University of Technology
Brisbane, Australia
Email: a.hodgson@murdoch.edu.au, f.maire@qut.edu.au

***Abstract*—This paper presents two algorithms to automate the detection of marine species in aerial imagery. An algorithm from an initial pilot study is presented in which morphology operations and colour analysis formed the basis of its working principle. A second approach is presented in which saturation channel and histogram-based shape profiling were used. We report on performance for both algorithms using datasets collected from an unmanned aerial system at an altitude of 1000ft. Early results have demonstrated recall values of 48.57% and 51.4%, and precision values of 4.01% and 4.97%.**

I. Introduction

Monitoring marine mammal populations is of great importance for the conservation and management of these species. Population estimates provide information about status which aids in developing conservation and management plans, while knowledge about migration patterns and critical habitat areas of species is important for maintaining the health of the marine ecosystems upon which they rely. Monitoring is typically performed by conducting regular surveys, which are used to detect changes to populations. This information, together with knowledge of the causes of any detected changes, is critical to ensure timely management measures can be prescribed to prevent further negative trends.

Regular status assessments of some marine mammal species used by the International Whaling Commission and the IUCN rely on minimum population estimates from aerial surveys. So too does the marine mammal monitoring that is required by law under the US Marine Mammal Protection Act of 1972. In Australia, regular surveys of dugongs ($Dugong\ dugon$) 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]. These data have been used to inform management by producing spatially-explicit population models as a proxy assessment of areas of high conservation value to dugongs [7]. Rapid assessments of the risks of human impact can then be conducted to determine the likely effectiveness of Marine Protected Areas [8].

Standard techniques for surveying marine mammals involve human observers flying along set flight paths in light aircraft, scanning sections of the ocean and recording sightings in real time. These aerial surveys allow coverage of large areas in a short period of time and have the advantages of (a) causing minimal disturbance to animal behaviour, and (b) providing an overhead view so that animals can be detected below the water surface. However, this method is expensive, involves risk to the human observers, and requires a team with very specialised skills for the identification of species. In addition, fatigue can result in missed sightings and misidentification of species. Therefore, automation of whole or part of this process would greatly benefit researchers by potentially improving the accuracy of aerial survey data. The use of still imaging payloads on-board either Unmanned Aerial Vehicles (UAVs) or manned aircraft for conducting aerial surveys is still in the early stages of development. The main impediment to using still imaging systems is the time required to manually review the tens of thousands of images that are acquired during a single survey.

Fig. 1. Representative image obtained from a aerial survey

In this paper we propose an approach for automatically detecting dugongs in aerial images captured using a customised payload onboard a UAV (Figure 1). We describe two novel pattern recognition algorithms that exploit specific features in

978-0-933957-40-4 ©2013 MTS

images. Our ongoing development of these automatic detection algorithms will contribute to the realistic use of new, improved, aerial survey techniques.

This paper is structured as follows. Section II reviews recent work on automated marine mammal detection. Section III describes the two approaches investigated. Section IV outlines the collection system and data used in the experiments. Section V presents the outcomes and analysis of data. Finally, section VI describes some of the lessons learnt and future work planned.

II. Background

The detection and monitoring of marine mammal species could potentially be conducted using a variety of different sensors, some complementary in the type of information they provide. For instance, active sensors (such as radar or sonar) may represent a potential approach to detect marine species [9][10]. They offer robustness against environmental conditions, however they tend to be complex, invasive and computationally intensive in terms of signal processing, and may be limited in their ability to discriminate between species.

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. However, this type of detection is often challenging due to the fact that marine mammals are diverse, and their vocalisations exhibit wide of variations in tone, frequency range and amplitudes, thus extensive knowledge of a species' vocal repertoire is needed in order to identify them. Additional knowledge required includes vocalisations rates and detection ranges [11] Imagery or visual observation techniques are continually being investigated but are limited by atmospheric conditions and greatly affected by illumination conditions [12], [13], [14]. Infrared and thermal imagery has the potential to improve upon visual observations by enabling nighttime detections [15], but again, may not have enough resolution needed to discern species. Despite its limitations, visual imagery or vision is an attractive solution given it offers a rich source of information. Cameras are inexpensive and have low power consumption which provides great advantages considering the inherent limitations in size, weight and power (SWaP) of many light aircraft or small/medium UAVs.

III. Detection approach

A. Morphological-based detection approach: Algorithm 1

In this section we describe the approach implemented in algorithm 1. This algorithm aims primarily to detect blobs in images. Colour values of pixels belonging to each blob are then analysed to further differentiate targets from background. The algorithm stages are:

1) Morphology: Our algorithm applies, as an initial stage, two fundamental operations in image processing called opening and closing, by combining two mathematical operations called erosion and dilation [16]. The aim of this step is to highlight elliptical features in the image that are suggestive of the marine mammal we aim to detect. This process is performed by applying an elliptical-shaped structuring element that resembles the shape of the marine mammal, to the grey level image. The size of this structuring element can be adjusted dynamically to account for the height of the camera above the surface.

2) Grey Level Thresholding: Thresholding is used to segment the image by setting those pixels whose intensity values are above a set point to a foreground value and all the remaining pixels to a background value. The aim is to create a binary image containing the areas (regions or blobs) of the elliptical shapes from step 1. We use adaptive thresholding techniques [17] to deal with changes in lighting conditions in the image, e.g., occurring as a result of a strong illumination gradient or shadows. This process is conducted as follows:

- Perform the mathematical convolution of the image with a suitable statistical operator, i.e. the mean or median. This convolution usually occurs in a local window of $n \times n$ pixels, where n can be 3, 5, 7.
- Subtract the original from the convolved image.
- Subtract from the above, a value C (where C is a positive constant typically in the range 0.02-0.06).
- Invert the thresholded image.

3) Blob detection: Blob detection refers to an approach aimed at detecting points and/or regions in the image that are either brighter or darker than the surrounding pixels. This technique creates a region counter that scans the image, for example, from left to right and from top to bottom. This is generally an iterative procedure that uses connected-component labelling to detect connected pixels in binary images. Connected pixels are then assigned to regions, by giving them common labels. Once regions have been labelled, the central moment of each region is extracted in order to find parameters such as area, centroid, perimeter, size, etc. From the set of detected blobs, we use criteria such as area and shape to filter those blobs that do not fit in an elliptical shape corresponding to the body shape of our target (dugongs and/or dolphins).

4) Colour analysis: Colour analysis is further used to improve the detection of targets by finding the colour intensity range by which dugongs will appear in images. This process is sub-divided into two steps:

i) Automatic colour description: This step characterises the colour of the image background in the HSV (hue, saturation, value) colour space and then defines threshold levels for dugongs based on this background colour description. This characterisation involves finding the average colour value for each channel in the HSV colour space, denoting each by h, s, v respectively. We found close correlation between the amount of background colour, e.g, blue, yellow, brown and the colour level by which each dugong appears in the image. This value is used to determine the appropriate colour level for thresholding described next in ii). Inspecting Figure 2 we can observe that levels of Hue, Saturation and Value for a set of images maintain a close relationship, specifically between the saturation and the average colour of the image background. Analysing several images such as the ones shown in Figure 3,

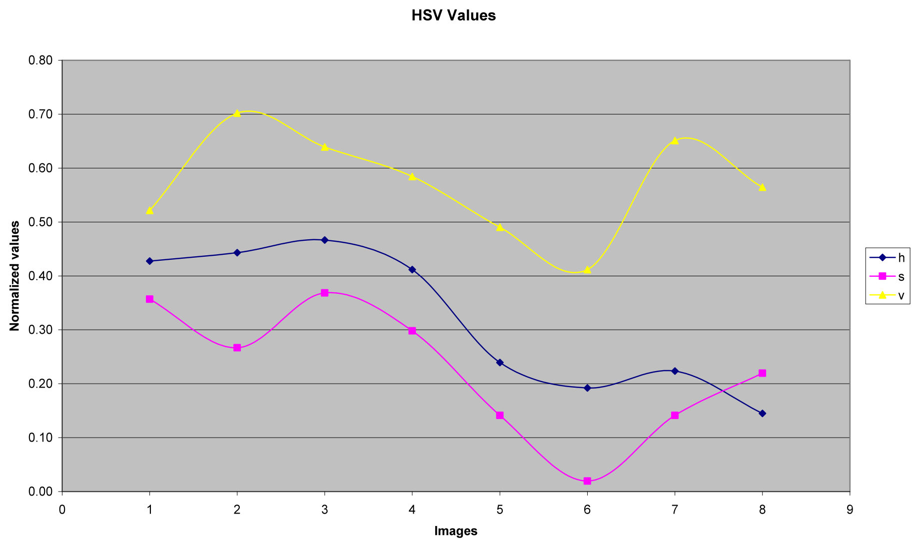


Fig. 2. Level of Hue, Saturation and Value for a set of images with similar properties to that in Figure 3

we found that images similar to Figure 3a lie in the range $h < 0.30$ and images similar to Figure 3b and 3c in $h > 0.30$. Using this information we then define a mapping function that associates HSV levels with a particular threshold. In this way, threshold values are adapted dynamically depending on the type of image.

ii) Blob colour filtering: Pixels within a blob are analysed to determine average colour level. Filtering is then performed combining step (i) with a data set made of several images of dugongs. From this data set, a colour value with a small tolerance (typically 5%) is then calculated and used to declare whether a blob is a dugong or not. For example, pixel values that are above a threshold (defined in (i)) are declared as dugongs, otherwise the blob is ignored.

B. Shape profiling on saturation channel: Algorithm 2

The approach followed in algorithm 2 differs from algorithm 1 by using the saturation channel from the HSV colour space, adding stages to deal with glint and analysing the shape profile of detected blobs. The algorithm consist of the following stages:

1) Saturation channel analysis: Similar to steps 1 and 2 in algorithm 1, however here we use the saturation channel (HSV colour space) of the image. Saturation can be defined as the ratio of colourfulness to brightness, generally referred to as the purity of the colour. The lower the saturation of a colour, the more grey and faded the colour will appear. A differentiating feature of algorithm 2 is that adaptive thresholding segmentation is applied to the HSV images to extract regions with low saturation values, rather than applying thresholding to the grey level image. The advantage is a more robust and straightforward detection given that dugongs appear grey or dark brown in aerial images

2) Glint removal: To further improve detection results, sun glint areas are removed before blobs are detected. Sun glint in images is created by reflection of direct solar radiation by waves on the surface of the water [17]. A glint brightness signature in the blue channel is characterised by pixel values greater than 240 (empirically derived from the 255 range). This bright region is often surrounded by a grey halo that has to be suppressed because it has the same colour range as that of dugongs. To remove sun glint areas, a mask is applied on the result of the adaptive thresholding performed on the saturation channel, followed by opening and closing morphology processes to remove or reduce remaining small blobs. This step has the advantage of reducing the number of candidate blobs that could be misclassified as dugongs.

3) Blob detection and feature extraction: Blob detection processing (as described in algorithm 1) is performed on the result of the glint removal step. Several features are computed for each detected blob: blob area as described in [18], length, width, and blob mean colour value in the HSV and RGB colour spaces. These features are then used to evaluate the match of the blob contour compared to a reference, using Hu moments [19].

4) Blobs filtering and features scoring: Blobs are filtered according to their area. A dugong is assumed to fit a specified area with a 20% margin relative to aircraft height (included in the image header). Feature scoring is based on computing the blob deviation from a reference. The reference is a dugong from a well-defined image taken at 500 ft altitude. We used the average of the six metrics to create a feature score: blob area, mean saturation colour deviations, surrounding ellipse length, width and eccentricity deviations and blob contour matching. The value of the feature score determines whether the blob is a positive detection or not.

5) Shape profiling: Shape profiling was inspired by Chow and Kaneko [17] with the distinction that our approach is based on histograms for fast processing. This step correlates a histogram of the blob shape with a reference histogram

a) b) c)

Fig. 3. Three examples of images with different colour properties used in the development of algorithm 1

to compute a profile score (Figure 4). The minimum and maximum values of the histogram indicate the body and tail sections of the dugong. The body section is then compared to an ellipse to produce an ellipsoid body error, and the tail length is compared to blob length to compute a tail to blob length ratio. If these two values are less than a threshold value, the blob is considered a positive detection.

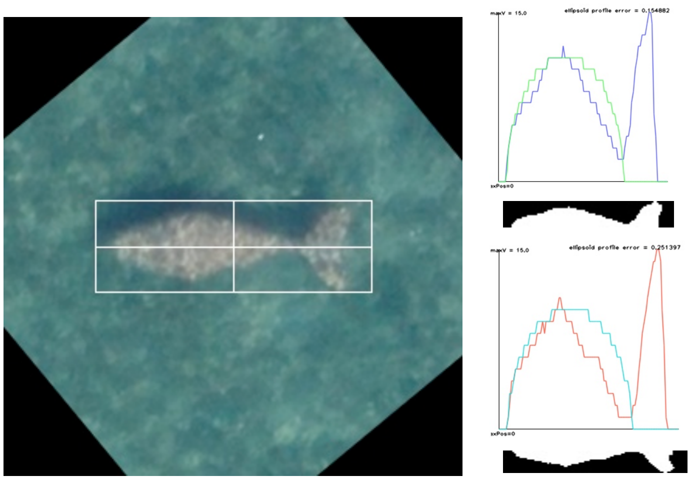
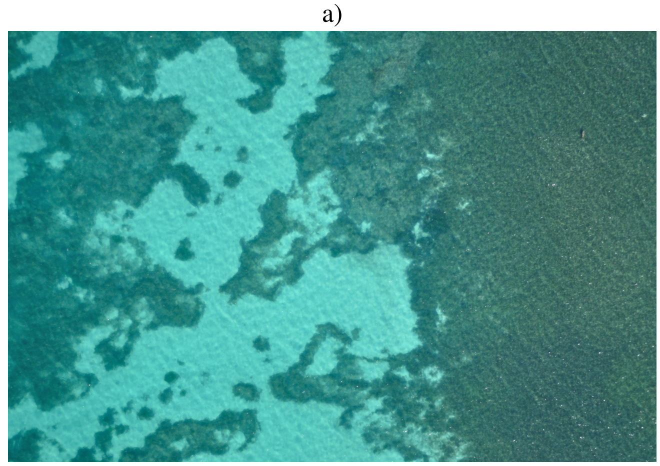


Fig. 4. Histogram profiling on dugong used as reference.

IV. Data collection and experiment setup

Our testing dataset consisted of images captured during a trial survey of dugongs using a $ScanEagle$ UAV operated by Insitu Pacific Pty Ltd in Shark Bay (Western Australia). 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 a series of trial flights, images were captured continuously along parallel line transects flown at three different altitudes: 500 ft., 750 ft. and 1000 ft. Transects were designed to cover different habitats (i.e., open water and seagrass banks in various turbidity levels) in an area where large numbers of dugongs were expected to occur. The flights occurred in a range of sea state conditions (Beaufort 0 to 5), and at various times of the day, creating varying levels of white caps and sun glitter in the images. We used a subset of 100 pictures captured at 1000 ft., including samples from a range of environmental conditions to evaluate both algorithms. Figure 5 shows a representative sample from this data set.

a)



b)

Fig. 5. Representative images taken at 1000 feet altitude showing the diversity in sea conditons.

V. Results

Using the data described in section IV which were collected in diverse environmental conditions, we evaluated both algorithms using two standard metrics in pattern recognition: recall and precision. Our results only allow us to draw conclusions within a limited scope given the lack of published benchmarks for this application. The key result from algorithm 1 was a high average false-positive value. Recall and precision for algorithm 1 was 48.57% and 4.01% respectively, with an average processing time per image of 120.9s (12Mpix images). The average false-positive obtained for this algorithm was 42.57 detections per image. Algorithm 2 achieved a recall of 51.4% and precision of 4.97% and a false-positive average per picture of 38.24. The average processing time per image was 9.14s.

These performance values indicated that both algorithms are able to return approximately half of the relevant results, however the proportion of retrieved detections that were truly dugongs was still low ($< 5\%$). Our continued development of this algorithm is therefore focused on reducing the number of false detections while continuing to improve the recall value. However, these algorithms represent a significant step towards automating the detection of marine mammals during aerial surveys using basic camera systems that can be adapted to both manned and unmanned aircraft.

Fig. 6. An example of the detection algorithm. Each detection is enclosed in red and a number is assigned to each. The algorithm also outputs a text file with this information once the processing is finished.

VI. Conclusion

This paper has introduced a valuable contribution to the automation of marine mammal surveys using aerial images. We have presented two algorithms aimed at detecting dugongs in images taken from a UAV. Our two versions of the algorithm are able to return relevant results at recall rates of 51.4% and 48.57%, which is a promising result. However, we now need to consider methods to reduce the algorithms' false-positive detections. An important feature of the algorithms presented here is the capability of cascading additional layers of processing within the algorithms to increase precision and robustness. This is a future avenue currently under investigation.

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] A. Grech, J. Sheppard, and H. Marsh, "Informing species conservation at multiple scales using data collected for marine mammal stock assessments," *PLoS ONE*, vol. 6, e17993, 2011.

[8] A. Grech and H. Marsh, "Rapid assessment of risks to a mobile marine mammal in an ecosystem-scale marine protected area," *Conservation Biology*, vol. 22, pp. 711–720, 2008.

[9] 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.

[10] 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.

[11] W. M. Zimmer, *Passive acoustic monitoring of cetaceans*. Cambridge, UK: Cambridge University Press, 2011.

[12] 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

[13] 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.

[14] 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.

[15] J. Churnside, L. Ostrovsky, and T. Veenstra, "Thermal footprints of whales," *Oceanography*, vol. 22, pp. 206–209, 2009.

[16] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*, 3rd ed. Upper Saddle River, NJ, USA: Addison-Wesley Longman Publishing Co., Inc., 2006.

[17] C. Chow and T. Kaneko, "Automatic boundary detection of the left ventricle from cineangiograms," *Computers and Biomedical Research*, vol. 5, no. 4, pp. 388 – 410, 1972. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0010480972900705

[18] G. R. Wilson, "Properties of contour codes," *IEE Proceedings - Vision, Image and Signal Processing*, vol. 144, no. 3, pp. 145–149, June 1997.

[19] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information theory*, vol. 8, no. 2, pp. 179–187, 1962.