Image analysis algorithms for automating the post-processing of cetacean sighting surveys data

B. Privat, G. Voleau, K. Lebart, Y. Petillot
School of Engineering and Physical Science
Heriot-Watt University
EH14 4AS Edinburgh, Scotland, UK
Email: K.Lebart@hw.ac.uk

R. Leaper
International Fund for Animal Welfare,
87-90 Albert Embankment,
London SE1 7UD, UK
Email: russell@ivyt.demon.co.uk

J. Gordon
Sea Mammal Research Unit
St Andrews University
St Andrews, Scotland, UK
Email: jg20@st-andrews.ac.uk

***Abstract*— Accurate measurements of the locations of surfacing cetaceans (whales, dolphins and porpoises) are important data for behavioral studies and sightings surveys. A system for tracking cetacean movements based on photogrammetric analysis of digital images, presented in [1], has been developed and tested at sea. This paper presents and discusses the use of image processing tools to partially automate the processing of the digital images thus produced, in order to obtain estimates of both the bearing and range of the sightings. It is hoped that these tools will be enablers of wider scale surveys, and they are expected to be deployed and field tested during trial in the spring/summer 2005.**

Fig. 1. Camera/video system. The down pointing digital camera is not shown here.

I. Introduction

Accurate measurements of the locations of surfacing cetaceans (whales, dolphins and porpoises) are important data for behavioral studies and sightings surveys. A system for tracking cetacean movements based on photogrammetric analysis of digital images, presented in [1], has been developed and tested at sea. Results of sea trials have demonstrated that the system is a practical tool for fine-scale tracking of cetacean movements and could also be used on line transect surveys.

Typically, the lightest system comprises a set of binoculars attached to a video camera covering the same field of view, and a downward looking digital camera to measure bearing relative to the ship's heading. When sighting a cetacean, the video recording is manually triggered, and the downward looking digital camera grabs a snapshot of reference patterns on the ship deck.

Radial distances from the ship to surfacing whales are calculated from the video images by measuring the angle of dip between the whale and the horizon. For this, scientists have to input through 3 mouse-clicks on a GUI two points on the line of horizon, and the position of the cetacean.

Bearings are measured from the still images of reference points on the ship. Here again, the reference points need to be identified manually through a GUI.

In this paper we present simple image processing techniques developed to be integrated within the system in order to automate the horizon detection and the bearing estimation, so as to minimize user intervention in the post processing of survey data.

For bearing estimation, a reference pattern in the shape of an isosceles triangle, pointing towards the front of the ship, needs to be installed in the field of view of the digital still camera. The camera axis is assumed to be perpendicular to the ship deck. The detection of the pattern is based on a Hough transform and a set of rules constraining the structure of the pattern. We present an analysis of the detection system's robustness to operational conditions: the effects of occlusions (the pattern is partially hidden), distractors (similar patterns or straight lines are added in the field of view), illumination changes (natural light), camera tilt and slant (changes the angles of the projected triangle), are discussed.

The horizon detection is challenging in terms of image processing, as the definition of the horizon line is very variable in the images. The methods presented are a diverse set of simple detectors, based on boundary detection, statistical region segmentation and prior knowledge.

The system is to be deployed in a wide area cetacean survey of the North Sea and adjacent waters in July 2005. Tests on operational data will ensue, and we expect the benefits of the automation in terms of expert time gain will be assessed then.

The paper first presents the state of the art in terms of methodologies to estimate the abundance of cetacean in section II, and motivates the need for automating the step requiring image analysis. Section III presents a solution for bearing estimation, while section IV tackles the automated detection of the horizon line required for range detection.

II. Cetacean Surveys: stakes and current methodologies

Estimates of abundance for whale, dolphin and porpoise (cetacean) populations are essential for conservation and often central to the debate about the threats they face including incidental takes in fisheries and commercial whaling. However, whales are not easy to count. The most common way of estimating whale abundance is to conduct a line transect survey based on visual observations. Such surveys can either be conducted from ships or aircraft. There is a large body of mathematical theory surrounding line-transect survey design and analysis (see for example [2]). The main data required are an accurate location of every cetacean sighted relative to the vessel or aircraft. These data are usually given in the form of a distance and bearing from the observer. Despite being critical to the final estimate, there has been relatively little progress in developing methods to measure either distance or bearing, and most surveys still rely on visual estimation. Some line-transect analyses are based on perpendicular distances from the trackline, in which case the error in the final abundance estimate will be inversely proportional to any bias in distance estimation. Other analysis methods are based on counting the number of whale cues (such as a surfacing) within an area (e.g. [3]) and for these methods the error in the final abundance estimate will be approximately inversely proportional to the square of any bias in distance estimation. The relationship between bias and variance in angle measurements and the final abundance estimate is more complex but equally important.

Some survey methods use observers searching with the naked eye while others require observers to search with binoculars. One advantage of searching with binoculars is that there is greater potential for instrumentation to assist with measurements of range and bearing. Currently, the most commonly used device to assist in the estimation of bearing is an angleboard that usually consists of a pointer which the observer lines up on the cetacean. However, cetacean surfacings are often of insufficient duration to allow this to be done while the animal is still visible, resulting in a large variance in the estimates. Although most surveys involve some experiments to attempt to measure the variance of angle estimates, these are usually done to static targets which are constantly visible and easier to sight. Thus, even these experiments do not represent all the sources of error to actual sightings of whales and underestimate the true variance. Analyses to measure variance in angle estimates based on data from two independent observer's estimates of the location of the same whale surfacing have been conducted ([4]) but these rely on determining that both observers saw the same whale.

This study attempts to improve the measurement of bearings to whales for observers using binoculars by measuring the direction relative to the vessel that the binoculars are pointing when a sighting is made. For the type of binoculars that are typically used for whale surveys with a field of view of less than 5o, it is sufficient to measure the direction of the binocular since observers naturally place the object of interest in the center. A simple technique for measuring the direction of binoculars based on images from a downward pointing digital camera was described in [1]. A camera was used in preference to an angle measurement device on a rigid pole and spindle to allow the observer freedom of movement to compensate for the motion of the vessel and to allow the observer to move about the deck to gain better vantage points. The camera took images of reference marks on the deck of the vessel that could be converted to relative bearings since the binoculars would be very close to horizontal when looking at whales at typical ranges of a few hundred meters. Although this system gave good results and proved accurate to the level at which the vessel heading could be determined, a disadvantage was the additional time taken to analyze the images.

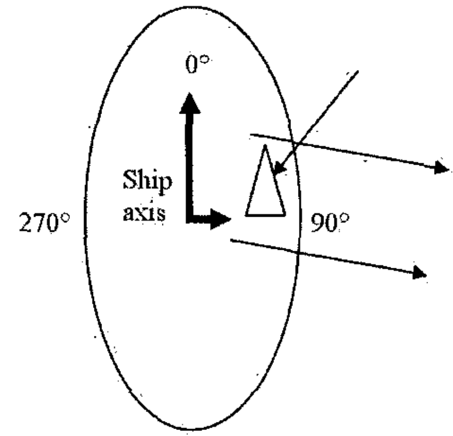


Fig. 2. Reference pattern and ship deck

In addition to measuring bearings to individual sightings, an automatic bearing measurement system has the potential to allow analysis of observer scanning patterns. This is important for modeling aspects of the detection process such as the probability of detection of a whale as a function of bearing. This will however require even more images to be analyzed greatly increasing the analysis load.

Leaper and Gordon [1] also describe a system for measuring range to whales based on images from video taken from a known height that include the horizon. The analysis process for these images relied on manually locating the horizon on the image. Similar image processing techniques to those described for bearings could also be used to automate the process of locating the horizon on an image.

The need to automate and streamline the analysis of images to measure range and bearing, and the more reliable and sophisticated population analysis that this would allow provide the impetus for the work to develop the image recognition software described.

III. Bearing Estimation

A. Principle

1) Reference pattern: Below the video camera used to record sightings is affixed a downward looking digital camera. This camera is triggered manually at each sighting, synchronously with the video recorder. For the purpose of bearing estimation, a pre-defined pattern of known and fixed orientation is put in the field of view of this camera. The pattern used is an isosceles triangle, oriented in the direction of the ship, as represented in fig 2. Typical images of the pattern would look like fig 3.

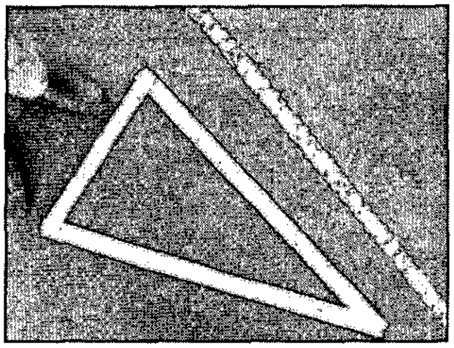


Fig. 3. Reference pattern

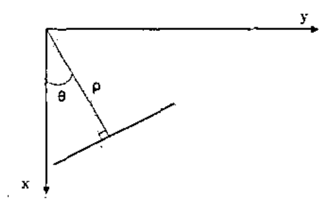


Fig. 4. Hough transform parameters

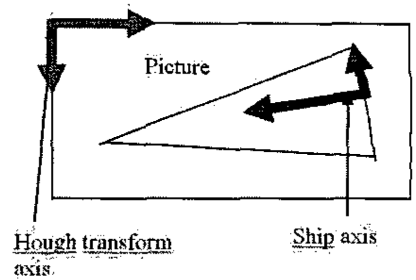


Fig. 6. Hough transform and Ship references

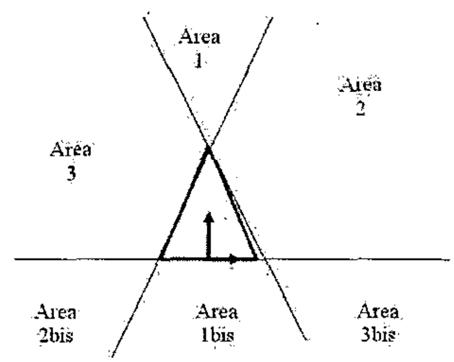


Fig. 7. The position of the origin of the image with respect to the triangle pattern is in one of the 6 areas outlined above

The bearing is estimated as the orientation of the camera in the horizontal plane. It can be extracted automatically from the image of the triangle: the orientation of the triangle is fixed and gives the heading of the ship. The orientation of the image compared to this triangle gives the orientation of the camera, and therefore the bearing of the sighting.

2) Hough Transform: The different stages of the analysis consist in detecting the triangle, extracting its constituting lines, and from these, deducing the orientation of the image. The main tool used to this effect is the Hough transform The Hough transform is a well known image processing technique, used to detect parametric curves within images. The simplest example of such curve is the straight line, that can be parameterized by two values. In this work with use the polar coordinate parameterization, that is, a straight line is uniquely represented by a pair (ρ, θ), as illustrated on fig 4.

In order to extract the constituting lines of the triangle, edges are first extracted from the image (fig. 5(a)).The binary image of the edges is then transformed into the Hough plane (fig. 5(b)). The points of highest intensity in the Hough plane (that is, the lines (ρ, θ) having the highest white pixel count) are determined by local maxima detections. They are potential candidates for being the sides of the triangle. The lines detected are overlaid with the original image in fig 5(c).

B. Bearing Estimation Algorithm

The algorithm estimates the orientation of the triangle within the picture, which in turns give the orientation of the camera with respect to the ship (cf. fig. 6). It scans each candidate line (ρ, θ) identified by the Hough transform and assesses whether it contributes to the triangle of the pattern. If three lines can be matched to the three sides of the triangle, then the orientation of the image is known.

When the picture is taken, the origin (top left corner) of the image will lie in one of the six areas determined by the sides of the triangle as illustrated on fig. 7. This yields in total 18 possible relative configurations of the triangle sides. These are each tested in turn by the algorithm, for each candidate line identified by the Hough transform, until a match is found. Note that the tests require the knowledge of both parameters, ρ and θ, for each lines.

C. Operational conditions study

In practice the system is to be deployed on a ship. The operational conditions are not fully controlled. The algorithm was informally tested on images simulating the typical operational conditions, as illustrated in fig 8.

Modifications to enhance the algorithm robustness have subsequently been implemented. We list below the aspects that are most likely to challenge the system, along with the steps taken to alleviate their effects:

(a) Edges (b) Hough Transform Plane (c) Image with detected lines

Fig. 5. Stages of the triangle detection

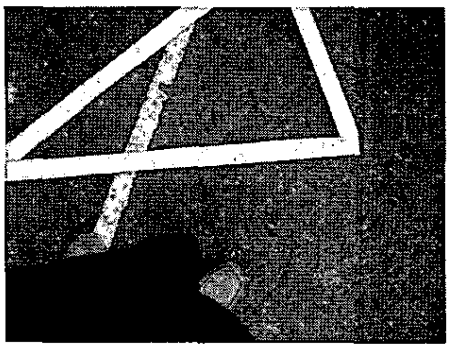


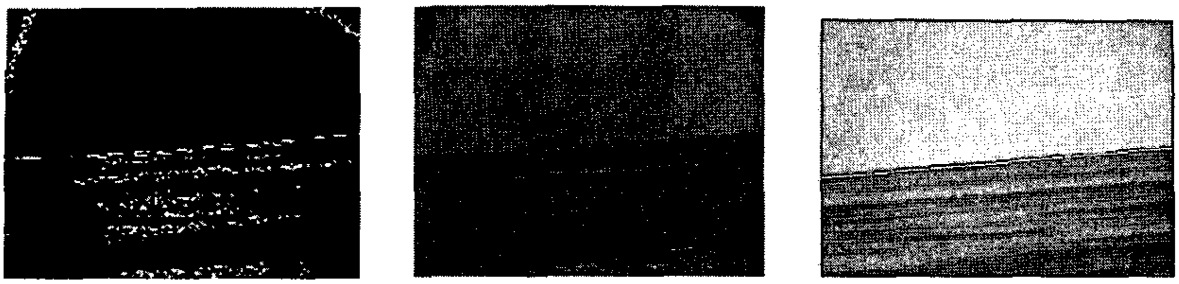
Fig. 8. Complex scene with a distractor (white line) and some occlusion and clutter

- Lighting conditions: they depend on the time of day, the weather, the presence or not of cast shadows, etc. One way to make the algorithm robust is to maximize the contrast between the pattern and the deck. The sensitivity of the edge detection can also be adapted automatically: the algorithm starts with a low sensitivity, and increases it if it fails to identify three lines. Beyond a certain sensitivity, the algorithm identifies it cannot give an answer and flags the problem.
- Distractors: amongst the various objects and parts of the ship structure that fall in the field of view, any straight line is a potential distractor, as it could be identified by the algorithm as one of the sides of the triangle. The strength of the contrast between the triangle and the background, as well as the fact that we actually look for a given calibrated structure, make the system particularly good at rejecting spurious line.
- Occlusions: various objects might partially cover the pattern (feet of the operator, clothes, life jackets, clutter from the equipment present on the deck). The algorithm is robust to partial occlusion, so long as it can detect the lines constituting the triangle. This can be done even if a number of parasite lines are taken into account after the Hough transform, as the constraint that the lines need to belong to the triangle enables to eliminate other candidate lines. The algorithm here again signals when it can no longer perform rather that give random outputs.
- 3D rotation: it is assumed that the down looking camera will be horizontal (and in practice, this is a good approximation). However, very small variations from the horizontal plane induce a perspective projection that changes the apparent angles of the triangle. Robustness can be gained towards these variations by loosening the calibration of the triangle. In this case a variation from the true value of the angles made by the sides of the triangle is tolerated. However, this is can have the detrimental effect of selecting spurious lines.

IV. Range estimation

In the system, range is estimated from video images. It currently requires the manual input of three points: two to define the horizon line, one to point at the mammal. The detection of sea mammals from these images is deemed particularly challenging, and out of the scope of this study. We concentrate on trying to extract the horizon line automatically, therefore potentially reducing the manual input time required from scientists during mission post-processing. The problem is quite challenging as variations in lighting, sea state, the presence of waves creating linear patterns, and the presence of relief (shore) in the horizon yield many different configurations. Here we outline solutions to first tackle simple cases where the horizon is clearly marked and there is no visible shore. Although they only represent a subset of the images, their automated processing can reduce significantly the need for tedious manual input from scientists.

Cases such as shown on fig. 9 seem quite straightforward. However, as shown on fig 10, standard techniques such as edge detection (fig 10(a)) or statistical segmentation based on grey levels (fig 10(b))require post-processing before the horizon can actually be detected. The option we select is to use the Hough transform on the edge (resp. region boundary) images. We also explore an alternative approach. It looks for interconnected horizontal segments, that split the image into maximally contrasting upper and lower halves. This is done by



(a) Edge detection

(b) K-means grey level segmentation

(c) Horizon line detected

Fig. 10. Horizon detection

Fig. 9. Horizon image

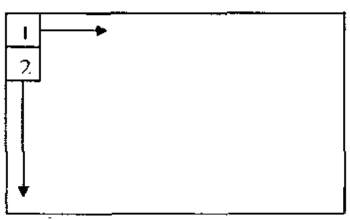


Fig. 11. Scan of the image by the double window

scanning a double window through the image, as represented in 11. The grey level distributions in each sub window (1 and 2) are compared using the following square distance measure d^2:

$$d^2 = (\mu_1 - \mu_2)^2/(\sigma_1.\sigma_2) \quad (1)$$

where μ_i, $i = 1,2$ is the mean grey level in each sub-window, and σ_i are the respective estimates of the grey level standard deviations in each sub windows. d^2 will therefore reach a maximum when the means of the distributions of each sub-windows are furthest apart, this separating distance being normalized by the standard deviations: while either of the sub-windows contains both sea and sky pixels, its standard deviation will be high, therefore bringing down the value of the distance.

In the first vertical scan, candidate positions where d^2 reaches a local maximum are logged. Subsequent columns are scanned in the vicinity of these positions. Fig 10(c) shows the horizon line detected by this method.

Preliminary studies on various images suggest that the fusion of different horizon detectors such as the ones suggested here will lead to a fairly reliable detection, as different detectors perform well on different images. This will be the object of further investigation.

V. Conclusion

In this paper we have presented a software solution for the partial automation of image analysis for cetacean sighting surveys. The system relies on standard image processing tools, and gives encouraging results on test data. It is hoped that it can be deployed and further tested during forthcoming surveys. The automation is provides is a potential enabler for larger scale surveys. Besides, it opens opportunities to further assess measurement precision and variability depending on various factors, such as the operator, or the range and bearing of sighting.

References

[1] R. Leaper and J. Gordon, *Application of photogrammetric methods for locating and tracking cetacean movements at sea*, J. Cetacean Res. Manage. 3(2):131-141, 2001.

[2] S.T. Buckland, D.R. Anderson, K.P. Burnham, and J.L. Laake, *Distance Sampling. Estimating abundance of biological populations*, Chapman and Hall, pp446, 1993.

[3] T. Schweder, H.J. Skaug, X.K. Dimakos, M. Langaas and N. Oien, *Abundance of northeastern Atlantic minke whales, estimates for 1989 and 1995*, Rep. Int. Whal. Commn. 47:453-484, 1997.

[4] T. Schweder, *Measurement error models for the Norwegian minke whale survey in 1995*, Rep. Int. Whal. Commn. 47:485-488, 1997.

[5] R. Gonzalez and R. Woods, *Digital Image Processing*, Pearson Education, Prentice Hall, 2nd Edition, 2001.