

Unsupervised Thresholding and Morphological Processing for Automatic Fin-outline Extraction in DARWIN (Digital Analysis and Recognition of Whale Images on a Network)

Scott Hale
Eckerd College
St. Petersburg, FL
halesa@eckerd.edu

Abstract

At least two software packages—DARWIN, Eckerd College [2], and FinScan, Texas A&M [4]—exist to facilitate the identification of cetaceans—whales, dolphins, porpoises—based upon the naturally occurring features along the edges of their dorsal fins. Both packages, however, require the user to manually trace the dorsal fin outline to provide an initial position. This process is time consuming and visually fatiguing. This research aims to provide an automated method (employing unsupervised thresholding and morphological processing techniques) to extract cetacean dorsal fin outlines from digital photographs thereby reducing manual user input. Ideally, the automatic outline generation will improve the overall user experience and improve the ability of the software to correctly identify cetaceans.

Keywords

morphological processing, thresholding, segmentation

1 Problem and Motivation

As cetaceans—whales, dolphins, porpoises—can be uniquely identified by the features of their dorsal fins, researchers frequently employ photo-identification techniques in studies of population, migration, social interaction, etc. Researchers construct a catalog of known individuals and classify fins based on the location of primary identifying damage features. The subsequent comparison of unknown fins to this catalog is time consuming and tedious. With the pervasive use of digital cameras in such research, the quantity of photographic data is increasing and the backlog of individual identification delays meaningful data analysis. Thus, automated fin comparison can potentially improve productivity. Automated methods to compare dorsal fin photographs exist and significantly reduce the number of photographs which must be examined manually to correctly identify an individual. However, these methods often require more work on the end user's part than simply looking through a catalog of photos. Tracing a fin outline accounts for a large portion of the time required for data input. Not only is the process time consuming, but also the nature of the task causes visual fatigue. In fact, some institutions still prefer the classic catalog approach to automated methods. Many institutions employ a digital catalog, but do not employ an automated-recognition package. In order to make these packages more usable and increase the end user experience, a

process to automatically extract a fin outline from a photograph is desirable.

2 Background and Related Work

2.1 Histogram Analysis Techniques for Segmentation by Thresholding

Segmentation is the process by which foreground objects are separated from the background in an image. This results in a binary image in which the foreground is one value and the background another. Thresholding is a popular image segmentation technique [8] that involves analysis of an image's histogram. Thresholding techniques are those that determine the threshold value t based on certain criteria. All pixels with values less than t are given one color while those greater than or equal to t are given another.

Thresholding methods are either global or local and point- or region-dependent [8]. Global thresholding algorithms choose one threshold for the entire image while local thresholding algorithms partition the image into subimages and select a threshold for each subimage. Point-dependent thresholding algorithms only analyze the gray level distribution of the image while region-dependent algorithms also consider the location of the pixels.

The p-title method is one of the earliest thresholding methods [3]. It determines a threshold t such that at least $(100-p)\%$ of the pixels in the image map into the

foreground, where p is the percentage of pixels the object(s) of interest is known to occupy a priori.

The mode method works well for images that are almost bi-modal, having distinct foreground objects and background. The threshold chosen corresponds to the valley of the histogram [5].

Where images are not bimodal, it is difficult to find a good valley in the histogram upon which to threshold. In this case, it is often possible to define a threshold at the "shoulder" of the histogram [6]. The histogram concavity analysis method is an extension of the mode method that considers points of concavity—valleys and shoulders—as threshold possibilities.

2.2 Binary Image Operations

A binary image is an image in which each pixel is one of two colors—usually white and black.

Standard Boolean operations—NOT, AND, OR, XOR—are defined between pixels of corresponding locations in two images of equal dimensions, with the exception of NOT, which requires only one image.

Open is a morphological technique so named for its tendency to separate joined objects—e.g. open the image. An open operation is comprised of one erosion followed by a dilation. An erosion turns edge pixels—black pixels with at least one white neighbor—white. Dilation performs the opposite operation: it turns white pixels bordering black pixels black unless doing so joins two disconnected regions of black pixels.

2.3 YIQ Color Space

YIQ color space separates an image's intensity information from its color information. Thus, intensity transformations may be easily realized on a color image. YIQ color space is closely related to the NTSC broadcasting standard [7] enabling color and monochrome television sets to display the same signal appropriately. The Y channel represents luminance or brightness. In moving an image from RGB color space to YIQ color space (Figure 1), the Y channel is formed by combining the RGB channels in

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Figure 1: An intensity image can be constructed by weighting each channel. The standard approach moves the RGB image to YIQ color space using the above equation. Y represents luminance/intensity, I in-phase, and Q quadrature.

portion to the human eye's sensitivity to each color [7].

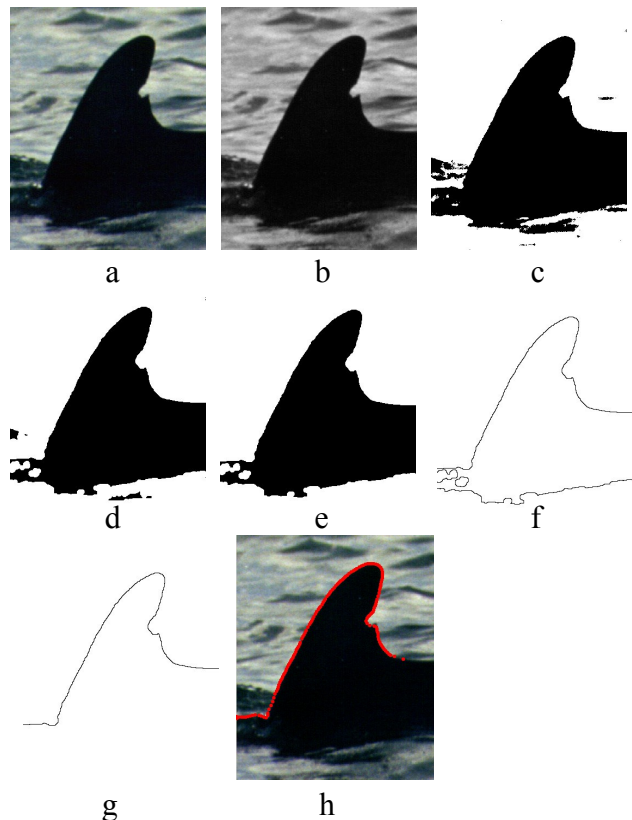


Figure 2: Intermediate steps for an automatic extraction of a fin's outline from a digital photo. (a) Color image, (b) Grayscale Image, (c) Result of unsupervised thresholding based on histogram analysis, (d) A series of morphological processes [open (erosion/dilation), erosion, AND with original binary image c], (e) Feature recognition to select the largest feature, (f) Outline produced by one erosion and XOR with e, (g) Feature recognition to select largest outline, (h) outline spaced/smoothed using active contours.

3 Uniqueness of the Approach

This research aims to determine what combination of image manipulation techniques reliably extracts dorsal fin outlines from digital photographs without user input. While the techniques are established, this aim is unique. The algorithm developed involves three primary stages—constructing a binary image, refining the binary image, and forming and using the outline yielded by the refined binary image.

The algorithm begins when the user imports a photograph of an unknown fin to compare to existing, known fins. The algorithm first seeks to establish a rough outline of the fin to define the constraint space and boundaries within which the canny edge detector and active contour operations may later operate. In

the first stage, the color image is converted to a grayscale image by moving the image to YIQ color space (Figure 2b) so that each pixel's intensity may be analyzed. The image is also downsampled for increased efficiency as only an approximation is necessary at this stage. A histogram of the grayscale image is constructed and analyzed. In order to keep the algorithm general and enable it to handle a wide array of images, only global point-dependent thresholding methods were surveyed. Point-dependent algorithms make no assumption about the location of the foreground object. Further, only global algorithms were considered as each image is assumed to consist of exactly one dorsal fin. The p-tile method [3] was surveyed and discarded in order to avoid fixing a percentage of the image that the dorsal fin must occupy, thereby limiting the domain of images upon which the algorithm would succeed. While optimal images are nearly bimodal and work well with the mode method [5], the histogram concavity analysis method [6] was found to work well on a larger spectrum of images and thus employed. The algorithm constructs a binary image (Figure 2c) in the second stage by thresholding the grayscale image at the threshold point chosen in stage one.

Morphological processes are applied to this binary image in the third stage to produce a cleaner outline (Figure 2d): often, thresholding will include dark portions of waves or shadows as part of the fin. The image is iteratively opened—erosion followed by dilation—to separate the fin from these small dark regions. An efficient implementation was found by adopting Cychosz's thinning algorithm [1]. The largest region of black pixels is then selected in the fourth stage (Figure 2e).

The outline is formed in stage five by eroding a copy of the image once and performing an exclusive or with the unchanged original binary image. This results in an outline one pixel in width (Figure 2f). The largest outline is then selected (Figure 2g) as smaller outlines often result from sun-glare spots on the fin and other abnormalities. In the next stage, the start and end of the fin outline are detected and walked to form a chain of coordinate pairs. The relative angles between the coordinate pairs are analyzed to determine the start and end of the dorsal fin as at this point the outline often includes the peduncle and other portions of the dolphin's body. Figure 3 shows an outline with the start and end points identified. In the final step (also performed for manually-traced fins), the algorithm plots the points

on the original color image and spaces/smooths the outline using active contours and a canny edge detector to reposition the outline along the true dorsal fin edge. (Figure 2h).

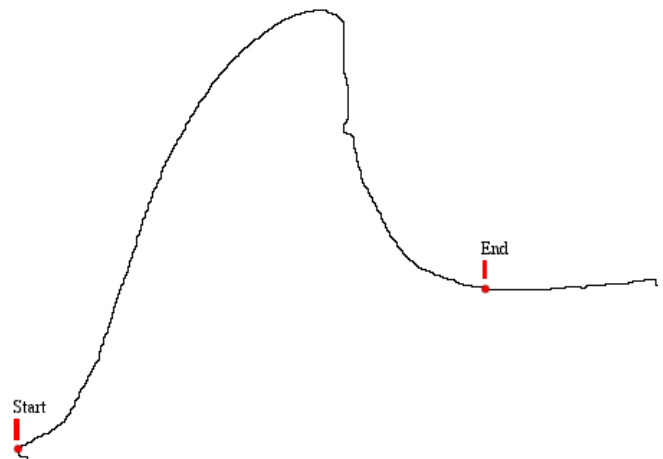


Figure 3: Outline with start and end points of the dorsal fin as determined by the angles between the points indicated.

4 Results

The fin extraction code was developed using a set of 36 images collected from field research. The test set of images included 302 images distinct from the 36 in the development set. These 302 images were collected as part of an active population study of bottlenose dolphins (*Tursiops truncatus*) in Boca Ciega Bay, FL. In the set, the program automatically produced a usable outline for 106 images (35.1%). Another 100 images (33.11%) required only slight modification to the automatically extracted fin outline. The autotrace algorithm was unable to produce an outline for 78 images (25.83%). Finally, the algorithm produced an unusable outline (usually selecting the wrong feature) for 18 images (5.96%).

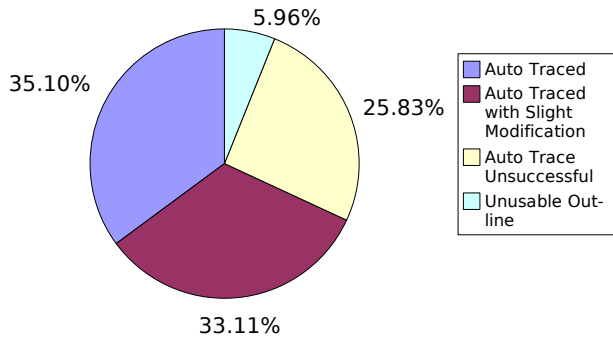


Figure 4: The results of a test set of 302 never-seen-before images.

Two-thirds of the images traced successfully—half with no modification whatsoever by the user. This significantly reduces the time and visual fatigue in fin outline extraction and greatly facilitates use of automated recognition packages. Given the large quantity of photographs to be compared, avoiding manually tracing two-thirds of the images greatly reduces end-user time and fatigue. Reactions expressed by marine scientists at the 16th Biennial Conference on the Biology of Marine Mammals suggest this research has the potential to improve user experience with automated recognition packages. In the event of failing to identify the outline, a user traces the outline as usual with no loss in time; while, in the case of a successful extraction of an outline, the user proceeds directly to matching the fin, bypassing the time-consuming and visually-fatiguing manual tracing process.

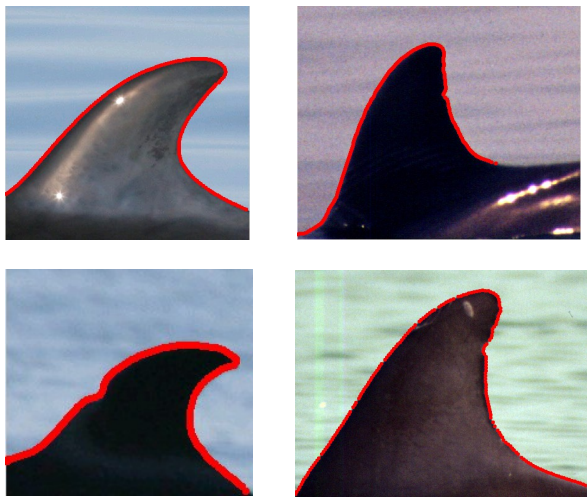


Figure 5: The outline extraction algorithm functions well for a wide variety of images.

5 Future

The algorithm recognized its failure to provide a usable outline for 78 images in the test set of 302. Some of the images were of poor quality in which sea-spray (Figure 6a) or water (Figure 6b) obscured the fin; however, others were images in which the standard intensity computation resulted in an image of low contrast between the fin and background. In these images, the fin and background regions had similar intensities despite visible differences in hue. The standard intensity image associated with an RGB image is formed by moving the image into YIQ color space with the Y channel representing luminance/intensity, the I channel in-phase, and the Q channel quadrature, using the equation in Figure 1. The next phase of this research focuses on methods to better utilize color information in the formation of an alternate intensity grayscale image. If the postulate that sea water has a higher intensity in the blue-green channels than does the fin proves true, the region with the lowest blue-green intensities should be the fin. If so, there exists a different matrix that when applied to the RGB image would produce an intensity image with a higher contrast between water and fin.



Figure 6: Sea spray (a) and wave/water obstruction (b) impede outline extraction.

6 Acknowledgments

This research was supported by the National Science Foundation under grant number DBI-0445126.

7 References

- [1] J. M. Cychoz, Efficient Binary Image Thinning Using Neighborhood Maps, *Graphic Gems, IV*. 465-473
- [2] DARWIN. Eckerd College. darwin.eckerd.edu
- [3] W. Doyle, Operation useful for similarity-invariant pattern recognition, *J. Assoc. Comput.* March 9, 1962, 259-267.
- [4] FinScan. Texas A&M University.

[5] J. M. S. Prewitt and M.L. Mendelsohn, The analysis of cell images, in *Ann. New York Acad. Sci.* Vol 1287, pp 1035-1053, New York Acad. Sci., New York, 1966.

[6] A. Rosenfeld and P. De La Torre, Histogram concavity analysis as an aid in threshold selection, *IEEE Trans. Systems Man Cybernet.* SMC-13, 1983, 231-235

[7] J.C. Russ, *The Image Processing Handbook*, 2nd ed. 1995

[8] P.K. Sahoo, et al., A Survey of Thresholding Techniques, *Computer Vision, Graphics, and Image Processing* 41. 1988, 233-260