Aquatic Mammals 2003, **29.1,** 117–123

Computer-assisted photo-identification of individual marine vertebrates: a multi-species system

G. R. Hillman[1], B. Würsig[2], G. A. Gailey[2], N. Kehtarnavaz[3], A. Drobyshevsky[1,4], B. N. Araabi[3], H. D. Tagare[5] and D. W. Weller[2,6]

[1]*Department of Pharmacology and Toxicology, University of Texas Medical Branch at Galveston, Galveston, Texas 77555–1031 USA*
[2]*Department of Marine Biology, Texas A&M University, Marine Mammal Research Program, 4700 Avenue U, Building 303, Galveston, Texas 77551 USA*
[3]*Department of Electrical Engineering, Texas A&M University, College Station, Texas 77843 USA*
[4]*Department of Neuroscience, Georgetown University, Research Building EP-04 3970, Reservoir Road NW, Washington, DC 20007–5927 USA*
[5]*Department of Diagnostic Radiology, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut, 06510 USA*
[6]*Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 8604 La Jolla Shores Drive, La Jolla, California 92037 USA*

Abstract

A computer program named 'Finscan' was developed for identifying individual marine animals by comparing new photographic images with a collection of previously identified images. The matching process was based on the pattern of nicks and notches commonly found along the trailing edge of the dorsal fin of many delphinid species. The program also allowed the inclusion of other user-defined descriptive features, such as leading-edge notches and truncated or irregular shapes. The output of the system was a presentation of images selected from the database, shown in order of similarity to a query image, so that the user could confirm the match. Two algorithms for representing notch patterns were tested and compared and the system was evaluated with dorsal fin images of several marine vertebrates, as well as fluke images for one species. Using a database of images that were previously identified by expert observers, the performance of the system was measured in terms of the number of incorrect matches that were offered before the correct match. Since in most cases the correct match was offered as the first or one of the first suggestions, the program substantially reduced the amount of effort required to perform photo-based matching.

Key words: photo-identification, computer, marine vertebrate, Finscan, cetacean, dolphin.

Introduction

Photo-identification of individual dolphins and whales has been an important part of research on the behaviour and biology of cetaceans since the early 1970s (Würsig & Würsig, 1977), and is now a common method used in many field studies (Würsig & Jefferson, 1990; Samuels & Tyack, 2000). Longitudinal studies incorporating individual recognition techniques have provided some of the most complete and detailed information on cetacean societies (Mann, 2000). As the duration of a particular study increases, so usually does the number of identified individuals. Several studies now span multiple decades, and massive photographic catalogues for a variety of species have been compiled. For example, catalogues of common bottlenose dolphins (*Tursiops truncatus*), humpbacked dolphins (*Sousa chinensis*), spinner dolphins (*Stenella longirostrus*), humpback whales (*Megaptera novaeangliae*), southern right whales (*Eubalaena australis*), bowhead whales (*Balaena mysticetus*) and sperm whales (*Physeter macrocephalus*) contain hundreds to thousands of recognized individuals (for summaries see Hammond *et al.* 1990; Mann *et al.* 2000).

As the number of identified individuals increases, the process of matching new images to those already catalogued becomes labour-intensive, and the probability of making errors is more likely. Each new photograph obtained in the field must be compared to all previously catalogued photographs to safeguard against false matches or missed

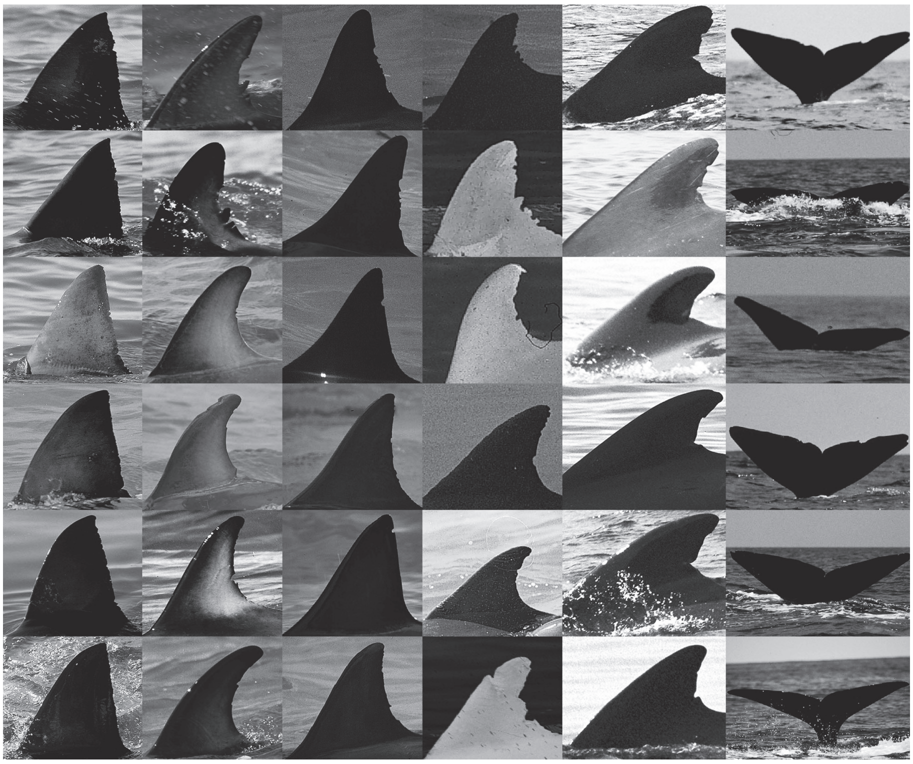

Figure 1. The top row shows a query image for each of six species. Below each query are shown five images retrieved by the computer system in order of similarity. The species, left to right, are: white shark, dusky dolphin, spinner dolphin, bottlenose dolphin, long-finned pilot whale, and sperm whale.

matches, a task requiring substantial time, effort, and skill. As subject catalogues grow into the hundreds and thousands, matching a single new photograph can sometimes take hours or days.

Several marine mammal computer-based photo-identification systems exist (Hiby & Lovell, 1990; Mizroch *et al*., 1990; Whitehead, 1990; Huele & Udo de Haes, 1998; Huele *et al*., 2000; Burnell & Shanahan, 2001; Hiby & Lovell, 2001) and each has helped to increase the efficiency in matching and individual recognition for the species for which it was designed. Although photo-identification methods based on commonly occurring nicks and notches found along the trailing-edge of the dorsal fin of numerous delphinid species have been used since the early 1970s, this line of study has benefited from only preliminary efforts to use computer-based image recognition technology (Kreho *et al*., 1997, 1999; Hillman *et al*., 1999, 2001; Araabi *et al*., 2000; Roberts *et al*., 2000; Gailey *et al*., 1999; Gailey, 2001).

This paper describes an accurate and efficient computer-assisted individual recognition system for some species of dolphins and whales (with a broader application to one shark species). The 'Finscan' system, developed using bottlenose dolphin dorsal fin images as the model, computes a description of the dorsal fin edge and compares it to a database of edges obtained from previously processed images. Finscan is user-interactive, allowing fin description, matching, and final confirmation of individual identification to be controlled by the computer operator. Although Finscan was developed for use with bottlenose dolphins, results presented here include a multi-species comparison of the system's performance for dorsal fin images of dusky dolphins (*Lagenorhynchus obscurus*), spinner dolphins, long-finned pilot whales (*Globicephala melas*) and white sharks (*Carcharodon carcharias*), as well as sperm whale fluke photographs (Fig. 1).

Materials and Methods

Finscan accommodated digital still, digital video frames, or 35-mm photographs, with the latter digitized via film scanner. A spatial resolution of 300 pixels on the longer edge of the image was standard, and the program adjusted images of varying sizes to this format. Images were 8-bit monochrome images in BMP, TIF, GIF, or JPG format. As new images were entered they were added to a database constructed using Microsoft Access data structures that contained the location of the image on disk, the identification of the animal (a temporary name until the matching process was completed), and the fin shape description. Finscan was written in Microsoft Visual Basic, with computationally intensive functions written in C.

When a new image was entered into the system, the 'Live Wire' algorithm (Mortensen & Barrett, 1998) was used to locate the boundary of the fin. This interactive algorithm was chosen rather than a fully automatic one because some images contain strong environmental artefacts such as glare, splashes, or other animals that obscure part of the fin edge of the target animal. In addition, the interactive nature of the live wire algorithm allowed users to supervise boundary detection as it proceeded, overriding any errors as they occurred, rather than reviewing and repairing unsatisfactory boundaries after automatic determination. The approximate tip of the fin, detected automatically by its curvature and location, was used as a landmark for one end of the trailing-edge curve and the approximate intersection of the dorsal fin with the animal's body served as the other end. These boundaries could be manually adjusted so that features occurring elsewhere on the fin (e.g., on the leading-edge) could also be included. The final selected distance was sampled or interpolated to provide 300 x–y coordinate pairs that described the overall boundary.

Once the boundary was extracted, two methods, *Curve Matching* and *String Matching*, were used to describe the shape of the fin for database comparison. Curve Matching used an estimate of the un-notched fin edge by smoothing the observed notched edge. A stiff deformable template, which bridges notches, was applied to the edge. The Euclidean distance of the notched edge from the smoothed model was then determined at each of 300 normalized points along the edge, resulting in a one-dimensional 'distance curve' that revealed a pattern of notches. The set of points that describe the fin edge, and the distance curve that describes the deviation of this edge from a smooth curve, are stored in the database as part of the information about each image. This pattern was compared to the corresponding patterns for fins in the database by simple point-by-point subtraction of the curve elements. The cumulative difference between the two curves indicated the degree of dissimilarity between the fins shown in the two images. Because the original scale and position of the image were not known precisely, the test curve was subjected to several stages of scale and offset variations for trial matches to each database curve, and the best fit was noted.

String Matching used the local first and second derivatives of the edge coordinates to compute a curvature function. Optimal smoothing of this function was determined by analysis of wavelet coefficients, removing small irregularities and pixellation noise while leaving the notch information (Kreho *et al.*, 1997; 1999). The smoothed curvature function was then reduced to a sequence of literal symbols (*a*, *b*) representing convexities and concavities, respectively. Strings of like symbols were merged into regions having measured heights and areas. A threshold that distinguished notches from noise was determined by unsupervised k-means clustering, and notches or prominences that were sub-threshold were considered to be non-curved and designated as class *c*. This process reduced the description of the fin to a brief sequence of *a*'s, *b*'s, and *c*'s, representing concave, convex, and flat regions, each with attributes that describe the width and depth of the edge segment. This string, like the notch-depth array, was recorded in the database. These attributed strings were compared using a variant of the Levenshtein method, in which a degree of dissimilarity is computed from differences in the string sequence, weighted by differences in the positions and depths of the curved regions. Comparison of these strings provided a quantitative measure of the similarity of any two such strings and was used for comparison of dorsal fin images. The String Matching method is described fully in a recent paper by Araabi and colleagues (Araabi *et al.*, 2000).

When a catalogue query was made, a new image was entered, the boundary was located, and the curve and string functions computed and stored. A search was initiated through the existing database using either the curve or the string matching method to compute similarity. Either method can be selected by a checkbox in the search window. In the present study both methods were used, separately, so that their results could be compared, but most users of the program would determine which method was more accurate for their data and would thereafter use that one. Database members (i.e., images of previously identified individuals) were ranked according to similarity to the query. Images of up to nine of the highest-ranked probable matches were displayed: when designing the system, we judged that display of multiple images in a three-by-three square arrangement made the images as small as could be seen clearly. The user can, however, choose to show fewer than nine images at a time, making the displayed images larger, and also can view any image at its full native resolution in a separate overlying window. Once the suggested images were displayed, the user confirmed the identification by a mouse-click on the correct match, or selected a button that displayed the next images in order of similarity. The user could scroll through the entire database, nine images at a time, if desired, before deciding on a match. Ultimately, the experienced user made the final identification, with the system providing preferred choices. The number of incorrect suggestions that ranked higher than the actual correct match determined the error rate of the system. In the experiment reported here the databases were chosen to contain matches to all test images, but in actual use the existence of a match is not known in advance, and the user would be able to view and compare a query image to the entire image database before declaring an animal as a 'new' individual.

The performance of the system was determined by the number of incorrect suggestions offered by the program as more similar to the query image than the true match. In a perfect performance, if a correct match was present, it would always be offered as the first choice. However, even if the true match was not ranked highest by the system, but was instead offered among the first few suggestions, the number of images that must be inspected was nevertheless greatly reduced, and thereby diminished the overall search effort required to find a match.

Results

System performance was tested by constructing small databases of animals that expert observers had previously matched manually and these identifications were assumed to be correct. Databases for the various species tested contained from 30 to 44 images, each representing a different individual. For inclusion in the database, images were judged subjectively to be of 'good' quality with respect to contrast and focus (i.e., to be of sufficient quality for individual identification) and to show animals with adequate markings for identification purposes. For each species database, 23 to 127 additional images were obtained that were known, also by expert identification, to match one of the database images, and the ranking of the correct match among those suggested by the system was recorded for each query.

Sets of images were obtained for bottlenose dolphins, dusky dolphins, spinner dolphins, long-finned pilot whales, sperm whales, and white sharks. The Finscan system was used in the same way for all species, except in the case of sperm whales, where flukes rather than dorsal fins were used. In this case, the left fluke tip was considered the 'top' of the fin, and the right fluke tip was considered the 'bottom' of the fin.

A group of observers experienced in photo-matching, working with a database of 500 images, estimated that in the present study the time required to locate the fin boundary and enter a new image into a collection was about 5 min using Finscan or when done manually (see protocol followed in Defran *et al*., 1990). The average time required to manually (without aid of a computer) find and confirm a match was about 19 min, while this process took about 5 min using Finscan. The computer-assisted system required only a few seconds to compute and rank the similarity of a query image to all database members, so most of the 5-min matching time was spent visually confirming the match. This finding illustrates the time saving advantage of Finscan, and we anticipate that the saving is even greater with larger databases.

Figure 2 shows the distributions of numbers of suggested matches offered by the system up to and including the correct identification. The Finscan system usually reduced the user's task to reviewing only a small number of images for comparison to a query image. In all cases, the median number of suggestions leading to a correct identification was quite small, indicating that the system was usually successful in presenting the correct match early in its ranked list of suggestions (Fig. 2). For the species presented herein, there was no consistent difference in performance between curve and string matching methods. In the case of pilot whales, the string method was noticeably more accurate. By inspection of the curve functions fitted to the dorsal fins, it was observed that some of the strongly curved fins of adult male pilot whales presented a problem to the curve matching method because the angles at the base of the fin, and in some cases the entire fin, were so concave that even the unbroken fin was treated as a notch. The string matching method was resistant to this error. The system performed somewhat less well on sperm whale flukes, a task for which it was not designed, than with bottlenose dolphins, whose fins have larger notches compared to their overall size.

Statistical analysis was performed to test the hypothesis that the Finscan program performed better than chance in identifying individuals. We made the assumption that if the program made matches solely by chance, then the mean number of suggested matches required to find a correct match would be half the size of the database from which matches were chosen. Therefore, for each species, we tested the hypothesis that the number of suggested matches actually required for a correct match was, over the number of trials for that species, less than half that population size in the database. The Wilcoxon signed-rank test was used to determine the P values. For each bar shown in

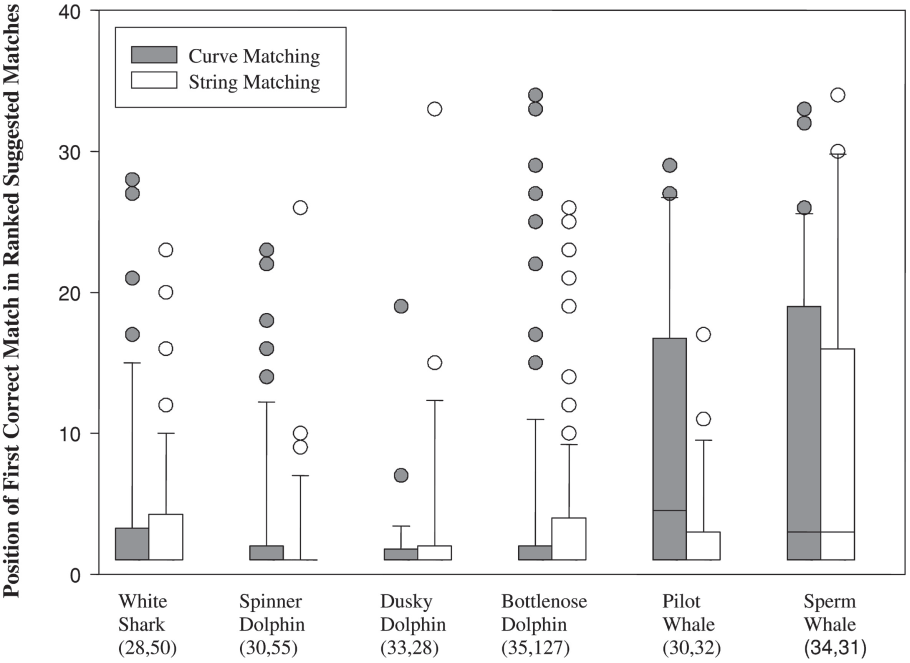

Figure 2. Figure shows the ranking of the known correct match among the suggested matches. A rank of 1 indicates that the first suggestion (the image scored by the computer as most similar to the query image) was the correct choice; a rank of 2 indicates that one image was judged to be more similar to the query than the correct match, etc. The crossbar (i.e., the median number of suggestions leading to a correct identification), the top of the bar, and the error bar represent, respectively, the 50th, 75th, and 90th percentile. If a crossbar is not visible, the 50th percentile corresponds to the bottom of the bar, indicating that in at least 50% of cases the first guess was correct. The circles represent individual outliers. Two matching methods were tested. The first number below each species name indicates the number of images, one per individual animal, in the database that was queried. The second number indicates how many images were used to query the database; each of these images was known to have a match within the database.

Figure 2, the number of suggested matches that were needed was less than half the population size at the $P<0.001$ significance level, except that for the sperm whale/string method bar ($P=0.0014$). These results indicate that the program is functioning to reduce the number of images that must be reviewed to a smaller number than would be obtained by a random search.

Discussion

Results of the test presented here demonstrate the versatility of the computer-assisted system; and generally excellent performance with dorsal fin images from dusky dolphins, spinner dolphins, bottlenose dolphins, and white sharks. Even with the small database sizes that were available to us for this study, Figure 2 indicates that for dorsal fins, in at least 50% of the images, a correct identification was made on the system's first suggested match, and for 75% of images a correct identification was made within the first three or four suggested matches. This performance is far better than chance, which would predict that on average the first correct match would occur halfway through the database, or on the 15th to 60th choice. However, performance could be improved for pilot and sperm whales if species-specific accommodations were made to the system. For example, observation of the fitted edges showed that the sharper curvature at the intersection of the fin with the body of the pilot whales caused error in some cases; adjustment of the curve fitting parameters could compensate for this. Likewise, the flexibility of sperm whale flukes causes the edge to be obscured in many cases: this problem could be addressed by treating each half of the fluke, from the central notch to the blade tip, as a separate edge, and matching only those edges that are fully visible.

An important concern not resolved here is the measurement of image quality (Friday *et al*., 2000; Gailey, 2001) upon which the performance of the system is dependent. The effective image quality is not determined by readily measured parameters such as contrast or focus, but is a complex signal-to-noise issue (including relative size and distinctiveness of identifying marks), resulting in the ability, or lack thereof, to perceive the shape of the dorsal fin. The inability to measure image quality interferes with a comparison such as the one presented here and presents a practical obstacle to the biologist who must decide whether or not to include available images in the searchable catalogue (for example, Gowans & Whitehead, 2001). However, the system is unaffected by rotation of the fin in the image plane and by the size of the fin in the image so long as the pattern of notches is not obscured by film grain or image pixellation. It is also robust to horizontal rotation out of the image plane (i.e., the dolphin moving obliquely toward or away from the camera), over a considerable range of angles, because the dorsal fin is thin enough to present the same apparent notch pattern even when seen at an oblique angle. It is less tolerant to elevated viewing angles, as the fin has finite thickness and the shape of the notches is distorted when seen from above. Experience has shown that while a quantitative measure of image quality remains elusive, the system works if the notch pattern can be seen in the image.

In addition, while the present system performs a linear search of small databases, larger image collections will require a faster search method. A kd-tree search method is under development, and will allow more efficient identification of the best matches, while allowing the search to be progressively expanded to include an arbitrary fraction of the database, as needed.

The Finscan system provides a content-indexed image database for use by field biologists. The tests performed here indicated that the system performs well enough to be genuinely helpful in the process of photo-identification by substantially reducing the time required to search for individual matches within pre-existing photo catalogues. Finscan is presently in use by several laboratories that are working with large collections of dorsal fin images. With input from these users, combined with further research and development by our team, we expect that within several years the performance of the present system will be vastly improved and used as a standard part of photo-identification studies. New software is also under development that will operate in a manner similar to Finscan but will perform matching based on patterns of spots, scars, and markings such as those on the flukes of gray and humpback whales.

Acknowledgments

We thank Michael C. Scholl of Dyer Island Shark Research Center for white shark images; T. Markowitz of Texas A&M University for dusky dolphin images; L. Karczmarski of Oceanic Society and Texas A&M University for spinner dolphin images; and A. Ottensmeyer of Dalhousie University for pilot whale images. We also thank two anonymous reviewers for suggestions that improved the manuscript. This research was supported by grants DBI 97-23000 and DBI 00-77661 from the National Science Foundation. Finscan can be downloaded at no cost from the following FTP site: ftp://utmb.wright.utmb.edu/

Literature Cited

Araabi, B. N., Kehtarnavaz, N., McKinney, T., Hillman, G. R. & Würsig, B. (2000) A string matching computer-assisted system for dolphin photo-identification. *Annals of Biomedical Engineering* **28,** 1269–1279.

Burnell, S. R. & Shanahan, D. (2001) A note on a prototype system for simple computer-assisted matching of individually identified southern right whales, *Eubalaena australis*. *Journal of Cetacean Research and Management (Special Issue)* **2,** 297–300.

Derfan, R. H., Shultz, G. M. & Weller, D. W. (1990) A technique for the photographic identification and cataloging of dorsal fins of the bottlenose dolphin (*Tursiops truncatus*). In: P. S. Hammond, S. A. Mizroch, & G. P. Donovan (eds.) Individual Recognition of Cetaceans: Use of Photo-Identification and other Techniques to Estimate Population Parameters, pp. 53–55 *International Whaling Commission Special Issue* **12,** Cambridge, England.

Friday, N., Smith, T. D., Stevick, P. T. & Allen, J. (2000) Measurement of photographic quality and animal distictiveness for the photographic identification of humpback whales. *Marine Mammal Science* **16,** 355–374.

Gailey, G., Würsig, B., Hillman, G., Kehtarnavaz, N., Drobyshevsky, A. & Weller, D. W. (1999) Computer-assisted photo-identification of bottlenose dolphins. Abstract, Thirteenth Biennial Conference on the Biology of Marine Mammals. Maui, HI., December.

Gailey, G. A. (2001) Computer systems for photo-identification and theodolite tracking of cetaceans. Unpublished M.Sc. thesis. Texas A&M University, College Station, TX. 125 pp.

Gowans, S. & Whitehead, H. (2001) Photographic identification of northern bottlenose whales (*Hyperoodon ampullatus*): Sources of heterogeneity from natural marks. *Marine Mammal Science* **17,** 76–93.

Hiby, L. & Lovell, P. (1990) Computer aided matching of natural markings: a prototype system for grey seals. In: P. S. Hammond, S. A. Mizroch, & G. P. Donovan (eds.) Individual Recognition of Cetaceans: Use of Photo-Identification and other Techniques to Estimate Population Parameters, pp. 57–61 *International Whaling Commission Special Issue* **12,** Cambridge, England.

Hiby, L. & Lovell, P. (2001) A note on an automated system for matching the callosity patterns on aerial photographs of southern right whales. *Journal of Cetacean Research and Management (Special Issue)* **2,** 291–295.

Hillman, G. R., Würsig, B., Gailey, G. A., Kehternavaz, N., Araabi, B. N., Weller, D. W., Mandava, S. & Tagare, H. D. (2001) Finscan: Computer-assisted photo-identification. Abstract, Fourteenth Biennial Conference on the Biology of Marine Mammals. Vancouver, B.C., November.

Hillman, G., Würsig, B., Kehtarnavaz, N., Weller, D., Kreho, A., Drobyshevskyy, O., Elder, K., Araabi, B., McKinney, T., Gailey, G. & Tagare, H. (1999) Individual identification of dolphins in field photographs: a computer-based system. Abstract, Thirteenth Biennial Conference on the Biology of Marine Mammals. Maui, HI., December.

Huele, R. & Udo de Haes, H. (1998) Identification of individual sperm whales by wavelet transform of the trailing edge of the flukes. *Marine Mammal Science* **14,** 143–145.

Huele, R., Udo de Haes, H. A., Ciano, J. N. & Gordon, J. (2000) Finding similar trailing edges in large collections of photographs of sperm whales. *Journal of Cetacean Research and Management* **2,** 173–176.

Kreho, A., Kehtarnavaz, N., Hillman, G. R., Würsig, B. & Weller, D. W. (1997) Computer assisted feature extraction for dolphin identification. *Proceedings of CISST'97 International Conference*, pp. 440–445. Las Vegas, NV.

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

Mann, J., Connor, R. C., Tyack, P. L., & Whitehead, H. (2000) *Cetacean Societies: Field Studies of Whales and Dolphins*. University of Chicago Press, Chicago.

Mann, J. (2000) Unraveling the dynamics of social life. In: J. Mann, R. C. Connor, P. L. Tyack, & H. Whitehead (eds.) *Cetacean Societies: Field Studies of Whales and Dolphins*, pp. 45–64 University of Chicago Press, Chicago

Mizroch, S. A., Beard, J. A. & Lynde, M. (1990) Computer assisted photo-identification of humpback whales. In: P. S. Hammond, S. A. Mizroch, & G. P. Donovan (eds.) Individual Recognition of Cetaceans: Use of Photo-Identification and other Techniques to Estimate Population Parameters, pp. 63–70 *International Whaling Commission Special Issue* **12,** Cambridge, England.

Mortensen, E. N. & Barrett, W. A. (1998) Interactive segmentation with intelligent scissors. *Graphical Models and Image Processing* **60,** 349–384.

Roberts, Z. W., Wilkin, D. J., Debure, K. R. & Reynolds, J. E. (2000) Efficient correction of differences in perspective, orientation and scale in DARWIN: digital analysis to recognize whale images on a network. Abstract, Thirteenth Biennial Conference on the Biology of Marine Mammals. Maui, HI., December.

Samuels, A. & Tyack, P. (2000) Flukeprints: A history of studying cetacean societies. In: J. Mann, R. C. Connor, P. L. Tyack, & H. Whitehead (eds.) *Cetacean Societies: Field Studies of Whales and Dolphins*, pp. 9–44. University of Chicago Press, Chicago.

Whitehead, H. (1990) Computer assisted individual identification of sperm whale flukes. In: P. S. Hammond, S. A. Mizroch, & G. P. Donovan (eds.) Individual Recognition of Cetaceans: Use of Photo-Identification and other Techniques to Estimate Population Parameters, pp. 71–77 *International Whaling Commission Special Issue* **12,** Cambridge, England.

Würsig, B. & Jefferson, T. A. (1990) Methods of photo-identification for small cetaceans. In: P. S. Hammond, S. A. Mizroch, & G. P. Donovan (eds.) Individual Recognition of Cetaceans: Use of Photo-Identification and other Techniques to Estimate Population Parameters, pp. 43–52 *International Whaling Commission Special Issue* **12,** Cambridge, England.

Würsig, B. & Würsig, M. (1977) The photographic determination of group size, composition, and stability of coastal porpoises (*Tursiops truncatus*). *Science* **198,** 755–776.