Finding similar trailing edges in large collections of photographs of sperm whales

R. Huele*, H.A. Udo de Haes*, J.N. Ciano+ and J. Gordon†

Contact e-mail: huele@cml.leidenuniv.nl

ABSTRACT

The North Atlantic and Mediterranean Sperm Whale Catalogue (NAMSC 1.0) contains images collected via the cooperative effort of several individuals studying sperm whales in the North Atlantic and Mediterranean. The collection offers an important opportunity to test matching algorithms as an aid to photo-identification of individual sperm whales. Of the 2,081 photographs in the catalogue, 1,929 were of sufficient quality for photo-identification. The trailing edge of the fluke, an identifying feature, was extracted by an interactive method. Subsequently, the trailing edge was represented in a normalised form by an affine transformation. Left and right halves were processed separately. Using different methods, 489 matching pairs of photographs were found. Based on these confirmed matches, the power of several measures of similarity was compared. The measure of similarity calculated by cross-correlating the continuous wavelet transforms of the extracted contours was found to perform best in practice. No conclusive matches between photographs from different geographic locations were found.

KEYWORDS: PHOTO-IDENTIFICATION; SPERM WHALE; EUROPE; ATLANTIC OCEAN; TECHNIQUES

INTRODUCTION

Sperm whale researchers working in the North Atlantic and Mediterranean have collected a considerable body of photographic material from sperm whales. The first edition of the North Atlantic and Mediterranean Sperm Whale Catalogue (NAMSC) was published on CD in 1999 by the International Fund for Animal Welfare (IFAW) and the Centre of Environmental Science at Leiden University (CML). Each submitted photograph was archived with the year and place of observation, and the name of the contributor. The copyright of the photographs remains with the contributors. The contour of the trailing edge of the fluke is used for identification of individual sperm whales (Whitehead and Gordon, 1986; Arnbom, 1987). Some evidence has been gathered (e.g. Dufault and Whitehead, 1995; Childerhouse *et al.*, 1996) to show that changes in the contour of the trailing edge are limited and allow matching observations up to at least a decade apart. The NAMSC will continue to evolve as new images are submitted but even now the collection offers an important opportunity to test contour matching procedures.

Relatively small collections (< 200 individuals) can be accessed using memory and visual inspection by an analyst, but some form of feature-based catalogue is needed to retrieve material from collections consisting of thousands of photographs. There are two general approaches to this: (1) by landmarking; (2) assigning a measure of similarity to a descriptor of the identifying features.

In landmarking, an observer assigns one feature out of a small set to a photographed individual. The photographs are subsequently stored in categories using the assigned feature as a key to the category. To match photographs, the observer only has to examine those photographs sharing the same feature stored in one category. Obviously, there is a trade-off between the resolution of the features and the size of the resulting categories. The method can be impressively effective, but may be susceptible to differences in personal interpretation of the identifying features. For example, two observers, or the same observer at two sessions, might assign different identifying features to the same individual and store photographs of one individual in different categories. Finding an incorrectly categorised item generally means re-examining the whole collection.

Methods based on mathematically derived measures of similarity between descriptors are insensitive to changes in human interpretation. However, while the human brain easily identifies two patterns as equal even if one of the patterns is transformed, designing a computer programme that equals this human ability has proved particularly challenging.

Both approaches have been applied in existing methods of automated photo-identification. Mizroch *et al.* (1990) uses a pure form of landmarking, while Hiby and Lovell (1990) apply similarity measure to mathematical descriptors. Whitehead (1990) uses a mixed form, applying a mathematical similarity measure to ascribed features.

Some progress has been made in identifying features from photos after affine transformation (Minh *et al.*, 1999; Tuytelaars and Van Gool, 1999), but these studies are limited to rigid features. The sperm whale fluke is flexible, and non-linear distortions, such as folding and bending, are superimposed on the linear transformations of roll, yaw and pitch, caused by the angle of the photograph. The fact that the identifying features are localised on the trailing edge, while the photographic projection is subject to non-linear distortion, makes the problem of automated matching of trailing edges a challenging one to address.

As a continuation of earlier work (Huele and Udo de Haes, 1998; Huele and Ciano, 1999) it was supposed that a measure of similarity between the wavelet transforms of two trailing edges would be less sensitive to non-linear distortions than a measure of similarity between the non-transformed trailing edges.

Ideally, the distribution of similarity measures of non-matching photographs would be completely disjunct from those of matching photographs. No such similarity measure was found and all measures produced false

* *Centre of Environmental Science (CML), Leiden University, PO Box 9518, 2300 RA Leiden, The Netherlands.*
+ *Florida Marine Research Institute, 7825 Baymeadows Way, Suite B200, Jacksonville, Florida 32256-7590, USA.*
† *Sea Mammal Research Unit, Gatty Marine Laboratory, University of St Andrews, Fife, KY16 8LB, Scotland, UK.*

positives and false negatives. False positives, i.e. photographs representing two different individuals but having a high similarity measure, are easily discovered by visual verification. False negatives, i.e. photographs representing one individual but having a low measure of similarity, can only be found by going through all the material. Both make it more difficult to find matching photographs, by increasing the bulk of material that has to be examined by eye. The performance of the measures of similarity was compared both on the fraction of false positives and on the fraction of false negatives.

MATERIAL AND METHODS

The NAMSC contains 2,081 photographs, digitised by staff at IFAW. Most photographs are black and white, although 61 colour photos are included. All photographs were digitised and stored as JPEG files, with a mean file size of 52kb. Of the 2,081 photographs, 152 could not be used for individual identification for reasons of quality, resolution, angle, or because the photo did not show a fluke at all. Flukes photographed at the beginning of the dive, with the trailing edge pointed downwards, were also discarded. See Table 1 for a breakdown into years and locations.

No totally observer independent method was found to extract the trailing edge from the photograph in a satisfactory manner. Conventional edge detection methods extracted not only the edge of the fluke but often also the meaningless edges of waves and clouds; therefore, a semi-automated technique was developed, in which the analyst identified the two tips of the fluke plus the central notch by mouse clicking. An algorithm was chosen, based on the modulus maximus method described by Mallat (Mallat and Hwang, 1992; Mallat and Zhong, 1992), to extract an approximation of the trailing edge. The analyst can either accept this contour as correct or can indicate which area of the curve does not correspond to the contour by dragging the mouse to indicate that particular area. The algorithm subsequently refines the approximation and again the analyst can either accept this or indicate a new area of interest. This process continues until the analyst is satisfied with the extraction. The contour is then stored as a structure consisting of the x and y coordinates of the two tips, the central notch and the contour.

The central notch on the trailing edge is the major singularity in the signal and was found to overwhelm the information for the rest of the trailing edge. For this reason, both halves were processed separately. The left half and right half were represented by interpolation as 64 numbers each and were normalised by the affine transformation as described by Tuytelaars and Van Gool (1999) and Schodts and Vleugels (2000).

For both the contours of the left half and the contours of the right half, four measures of similarity were calculated:

(1) linear correlation between the contours;
(2) the maximum value of the cross-correlation between the contours, based on the argument that the contours might be shifted in phase due to noise in the extraction method;
(3) linear correlation of the wavelet transforms of the contour; and
(4) the maximum value of the cross-correlation of the wavelet transforms of the contour, based on the assumption that this measure would be relatively insensitive to linear and non-linear distortions of the contour.

The separate measures were combined in two different ways to calculate a measure of similarity for the total contour: (1) arguing from similarity, the measure was defined as *r1_total* = *r_left* * *r_right*; and (2) arguing from dissimilarity, the measure was defined as *r2_total = 1-(1-r_left) * (1-r_right)*.

As wavelet transformation, the bi-orthogonal wavelet 6.8 was chosen after some preliminary tests. The transformation results in a multi-resolution, two-dimensional representation of the trailing edge. Traditionally, the transform is depicted as an image, where the colours or shades of grey indicate the values (see Fig. 1).

Table 1

The number of photographs in NAMSC, per year and location, number of photographs suitable for matching, number of matches and number of identifiable individuals, per location.

	Andenes	Azores	Canaries	Dominica	Grenada	Jamaica	Mediterranean	Panama	Shetland	Greenland	Total
1984	-	1	-	15	-	-	-	-	-	-	16
1985	-	1	-	-	-	-	-	-	-	-	1
1986	-	-	-	-	-	-	-	-	-	-	0
1987	7	24	-	-	-	-	-	-	-	-	31
1988	14	156	-	-	-	-	-	-	-	-	170
1989	30	116	-	-	-	-	-	-	-	2	148
1990	48	133	-	-	-	-	-	-	-	8	189
1991	24	178	-	-	-	-	-	-	-	-	202
1992	25	-	-	-	-	-	-	-	-	-	25
1993	31	123	21	-	-	10	-	10	-	-	195
1994	30	126	-	-	-	-	39	-	-	-	195
1995	18	172	7	73	8	-	1	-	-	-	279
1996	53	161	3	26	-	-	2	-	-	-	245
1997	79	116	-	4	-	-	1	-	-	-	200
1998	65	113	-	-	-	-	3	-	-	-	181
1999	-	-	-	-	-	-	-	-	1	-	1
Undated	-	1	2	-	-	-	-	-	-	-	3
Total	424	1,421	33	118	8	10	46	10	1	10	2,081
Acceptable	378	1,322	33	117	8	10	41	10	0	10	1,929
Matches	80	297	45	17	2	0	47	0	0	1	489
Individuals	318	1,134	15	105	8	10	28	10	1	9	1,638

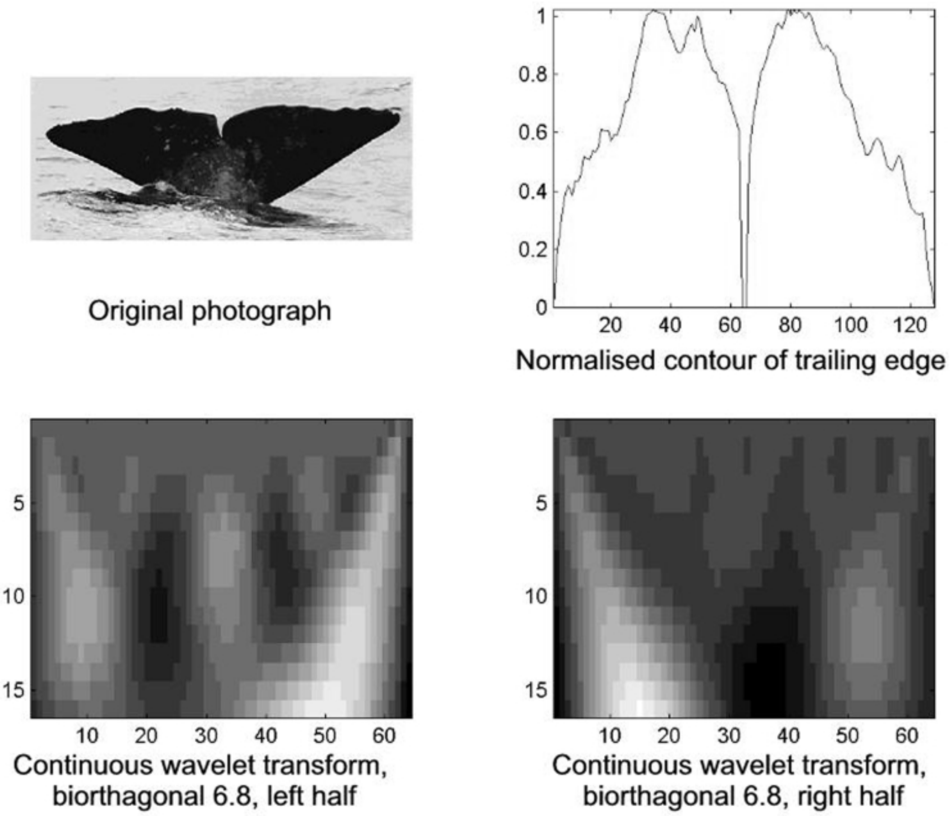

Fig. 1. The transformation of the trailing edge.

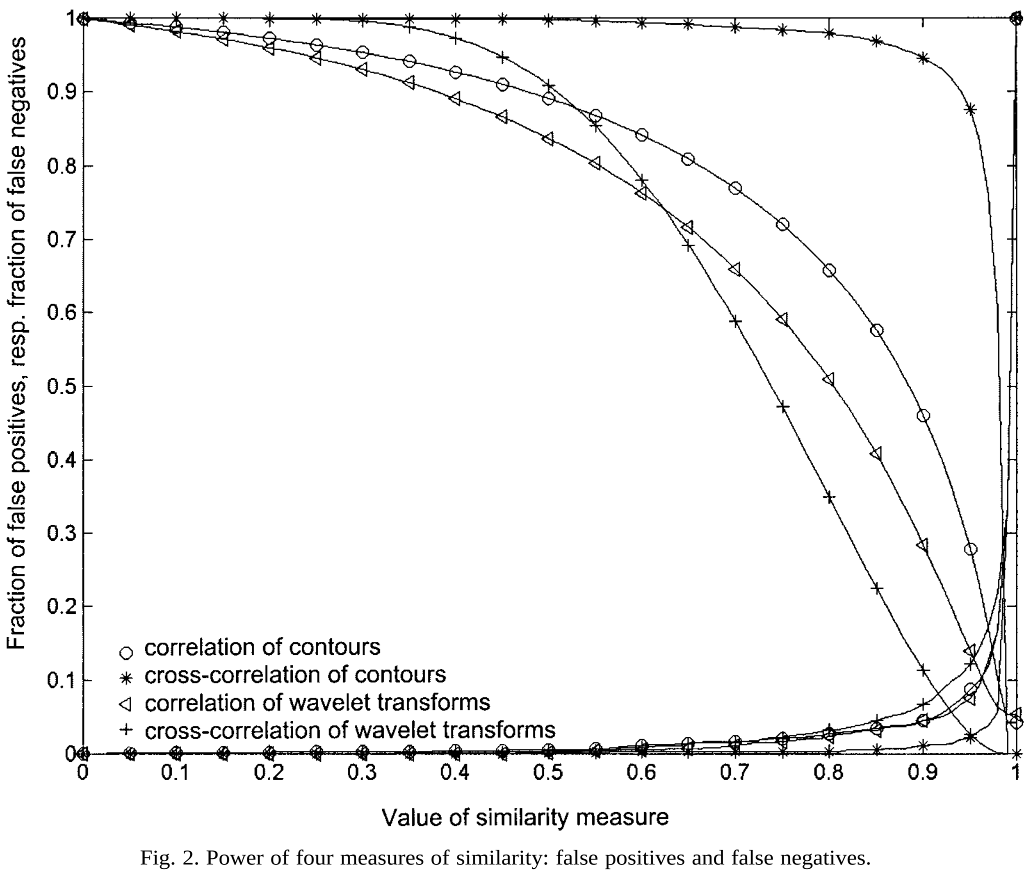

Fig. 2. Power of four measures of similarity: false positives and false negatives.

RESULTS

For the 1,927 photographs that were considered, 489 matches were found, counting one set of different photographs of one individually identified sperm whale as one match. Matches were confirmed by eye by at least two independent observers. No similarity measure, combined with a threshold value, separated all matching pairs from the non-matching pairs. In all cases, the measure of the total defined by *r_total = 1- (1-r_left) * (1-r_right)* was found to perform better than the measure calculated by direct multiplication. Likewise, the combined measure was found to perform better than the similarity measures of the separate halves.

Comparing the different measures of similarity, it was found that cross-correlating the contours produced the least false negatives (Fig. 2). However, this method caused many false positives, which made it difficult to find a true match in a sub-collection of sets having high measures of similarity. Cross-correlating the wavelet transforms led to a slightly higher number of false negatives, but discriminated more strongly on false positives. As sets of photographs always had to be examined by eye to be confirmed, the latter method proved more practical for choosing the 'nearest neighbours' as likely candidates.

Digitising and archiving the photographs is relatively time consuming. Once the photographs are in digital format, it takes about one hour to extract the contours of 100 photographs. Cross-correlating all 2,081 photographs against each other took eight hours of calculation time per level, on a 133Mhz Pentium machine with 128-Mb internal memory. It is estimated that scanning the 2,081 photographs and extracting the contours required about 80 hours. Calculating the wavelet transforms takes a few minutes. Calculating the cross-correlations took another seven hours of computer time on this machine. After the preparations, it took approximately six seconds to find the nine most likely candidates for matching to a given photograph.

DISCUSSION AND CONCLUSION

User input is needed to distinguish the trailing edge from other edges in the photograph. Photographs of high quality and taken according to a strict protocol based on Arbom's 'measure of photoquality' (Arnbom, 1987) could generally be analysed without any correcting input from the analyst.

Due to the loss of information caused by the projection, the lack of independent confirmation of found similarities between contours, the possibility of contours changing over time and the existence of featureless trailing edges, it will be impossible to design a totally automated method to match different photographs of one individual.

Cross-correlating the wavelet transforms performed as measure of similarity is best in practice, but it seems to be quite possible that a better performing measure of similarity can be found.

Regrettably, no ordinal ordering was found, making it impossible to prove the absence of a matching photograph in the collection.

The set of nearest neighbours, presented in response to a given photograph, occasionally contained mutually matching photographs, although neither of these photographs matched the query image. Further analysis of this effect may be useful for the improvement of the algorithm.

Only a small set of photographs was in colour and this extra information was not used. However, experiments showed that the colour information is very useful in the extraction of the contour, and therefore the use of colour photographs is recommended.

Digital collections greatly facilitate the distribution of material and the possibilities of following individual sperm whales over longer periods and wider areas. The NAMSC has shown that the material can be shared while the copyright, and possible further details of the observation, remain with the contributor. The possibilities of setting up a world catalogue of sperm whale photographs deserve further study.

ACKNOWLEDGEMENTS

The NAMSC is the collective product of Alexandros Frantzis; Alnitak Project, Spain; Andenes Cetacean Research Unit, Norway; Centro Interdisciplinare Di Bioacustica, Italy; CML, Leiden University, Netherlands; Dalhousie University, Canada; Department of Oceanography and Fisheries, Azores; Groupe de Recherche sur les Cetaces, France; Gulf-Cet, USA; International Fund for Animal Welfare; Michel André; Shetland Sea Mammal Group, Scotland; Sea Watch, Tethys Research Institute, Italy; and Whale Watch Azores. Tim Lewis and Lisa Steiner scanned the material for the NAMSC and designed the database and interface. Hans van den Berg highlighted the rationale of using wavelets. Bert Schodts and Bart Vleugels suggested the affine transformation. Peter van der Gulik highlighted the possibilities of the method. The two anonymous reviewers stimulated the more thorough analysis presented here.

REFERENCES

Arnbom, T. 1987. Individual identification of sperm whales. *Rep. int. Whal. Commn* 37:201-4.

Childerhouse, S.J., Dawson, S.M. and Slooten, E. 1996. Stability of fluke marks used in individual photo-identification of sperm whales at Kaikoura, New Zealand. *Mar. Mammal Sci.* 12:447-51.

Dufault, S. and Whitehead, H. 1995. An assessment of changes with time in the marking patterns used for photoidentification of individual sperm whales, *Physeter macrocephalus*. *Mar. Mammal Sci.* 11(3):335-43.

Hiby, A.R. and Lovell, P. 1990. Computer aided matching of natural markings: a prototype system for grey seals. *Rep. int. Whal. Commn* (special issue) 12:57-61.

Huele, R. and Ciano, J.N. 1999. Using wavelet transforms to match photographs of individual sperm whales identified by the contour of the trailing edge of the fluke. pp. 753-60. *In:* P. Dionysius, A.A. Huijsmans and W.M. Smeulders (eds.) *Visual Information and Information Systems.* Lecture Notes in Computer Science No. 1614. Springer Verlag, Berlin.

Huele, R. and Udo de Haes, H.A. 1998. Identification of individual sperm whales by wavelet transform of the trailing edge of the flukes. *Mar. Mammal Sci.* 14(1):143-5.

Mallat, S. and Hwang, W.L. 1992. Singularity detection and processing of wavelets. *IEEE Trans. Inf. Theory* 8(2):617-43.

Mallat, S. and Zhong, S.F. 1992. Characterization of signals from multiscale edges. *IEEE Trans. Pattern Analysis Machine Intell.* 14(7):710-32.

Minh, D., Ayer, S. and Martin Vetterli, M. 1999. Invariant image retrieval using wavelet maxima moment. pp. 451-8. *In:* P. Dionysius, A. Huijsmans and W.M. Smeulders (eds.) *Visual Information and Information Systems.* Lecture Notes in Computer Science No. 1614. Springer Verlag, Berlin.

Mizroch, S.A., Beard, J. and Lynde, M. 1990. Computer assisted photo-identification of humpback whales. *Rep. int. Whal. Commn* (special issue) 12:63-70.

Schodts, B. and Vleugels, B. 2000. Design for an image recognition algorithm, applied to whale flukes. Paper available from Vesaliusstraat 13, B-3000 Leuven, Belgium. [E-mail: *groept@groept.be*]. [In Dutch].

Tuytelaars, T. and Van Gool, L. 1999. Content-based image retrieval based on local affinely invariant regions. pp. 493-500. *In:* P. Dionysius, A. Huijsmans and W.M. Smeulders (eds.) *Visual Information and Information Systems.* Lecture Notes in Computer Science No. 1614. Springer Verlag, Berlin.

Whitehead, H. 1990. Computer assisted individual identification of sperm whale flukes. *Rep. int. Whal. Commn* (special issue) 12:71-7.

Whitehead, H. and Gordon, J. 1986. Methods of obtaining data for assessing and modelling sperm whale populations which do not depend on catches. *Rep. int. Whal. Commn* (special issue) 8:149-65.