TOWARDS COMPUTER-ASSISTED PHOTO-IDENTIFICATION OF HUMPBACK WHALES

Elena Ranguelova, Mark Huiskes, Eric J. Pauwels

CWI, Kruislaan 413, 1098SJ Amsterdam, the Netherlands

ABSTRACT

This paper describes current work on a photo-id system for humpback whales. Individuals of this species can be uniquely identified by the light and dark pigmentation patches on their tails (flukes). We developed an interface that assists the user in segmenting the animal's tail from the sea and fitting an affine invariant coordinate grid to it. A numerical feature vector capturing the patch-distribution with respect to the grid is then automatically extracted and used to match the individual against a database of similarly processed images.

1. INTRODUCTION

Individual identification of cetaceans (the marine mammals, i.e. whales, dolphins and porpoises) is of great interest to marine biologists. Identification plays an important role in their long-term studies of the population and behavioural patterns of the mammals [7, 5, 1]. Due to practical and legal issues associated with tagging the mammals, the method of photo-identification offers a valuable non-invasive alternative. This approach hinges on the uniqueness of the natural markings which can be captured by photographing the dorsal fins or flukes (i.e. tails). Marine biologists discovered more than 30 years ago that humpback whales exhibit sufficient variation in their natural markings to allow the identification of individuals based on images of their flukes. The markings are the white/light grey patches as well as blotches and long-lasting scars (some examples are shown on the next pages).

As the photographic collections grew, so did the need for more efficient retrieval methods that would allow a researcher to quickly compare and match new photographs against a database of registered individuals. There are several approaches for photo- identification found in the literature. In [7] manually generated code is used, based on a set of 38 generic fluke patterns and which takes into account the shape of the central notch and the location of blotches/scars (relative to a subregional division of a standardised fluke). After the coding the resulting descriptors are subsequently used for matching. Similarly, WhaleNet [9] is a graphical user interface (GUI) which allows the user to narrow down the search for matches by visually selecting one of 18 fluke types. Araabi extends a curve-matching technique, originally developed for the identification of bottlenose dolphins [2], for the encoding of the fluke's trailing edge of humpback whales [1].

The approach proposed in this paper comprises two main steps, namely the extraction of the fluke region and patches, and the actual matching. Similar ideas have been discussed in [5], but our approach differs in the methodology used. Kehtarnavaz et al. [5] introduced an interactive live-wire algorithm for the fluke extraction while we favour the use of morphological segmentation tools. And while [5] introduces affine moment invariants as features, we avoid the use of (numerically unstable) high order integrals by constructing an affine invariant grid that is automatically fitted to the fluke, dividing it into relatively small, but invariant regions. Next, each region is characterised by the relative contribution of dark and light patches. This maps the visual information into a numerical feature vector which can then be compared to the feature vectors obtained from other images.

2. EXTRACTION OF TAIL AND FLUKE PATCHES

The proposed matching approach requires an accurate segmentation of the fluke in white and dark patches. Because the photographic material is typically quite challenging (small colour differences between animal and background, confounding factors such as water splash, highlights on wet surfaces, etc.) automatic segmentation is unable to deliver the accuracy required for photo-identification. Thus, we have opted for semi-automatic segmentation based on a *marker-controlled watershed algorithm* [3].

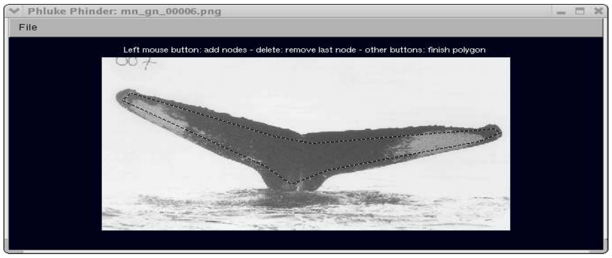
The watershed transformation is a powerful and well- established mathematical morphology tool for image segmentation which has been used in many applications [3]. Any grey-level image can be considered as a topographical surface. Flooding this surface from its minima while preventing the merging of water coming from difference sources, will result in a partitioning of the image into *catchment basins* associated with each minimum. The boundaries between the catchment basins are the watershed lines. If we apply this transformation to the gradient of an image, we should obtain catchment basins corresponding to homogeneous grey-level regions. It is well-known however, that the transform tends to produce an over-segmentation due to the local variations in the gradient. A *marker-controlled* transformation is a solution to this problem: The gradient image is modified via morphological reconstruction [3] in order to keep only the most significant gradient edges in the areas of interest between the markers.

The tail extraction process is initialised by the user, who specifies a rough initial contour (marker) within the tail (Fig.1 (a)). The watershed transformation is then applied to the modified gradient and automatically produces an estimated boundary contour for the tail. In many cases this first estimate will be acceptable for further processing. However, whenever needed, the segmentation interface allows the user to fine-tune the result by interactively introducing set of additional *positive* and *negative* markers. The positive markers indicate regions which should belong to the tail, but were not included in the segmentation while the negative markers indicate areas within the tail which were included in error. In the current implementation of the interface, it is possible to pre-process the image using median filtering and also to smooth the contour using morphological operators. Finally, the noise- and error-prone region at the basis of the tail (due to wave occlusion, water splash, etc.) is removed by clipping the contour (Fig.1 (b)).

0-7803-8554-3/04/$20.00 ©2004 IEEE.

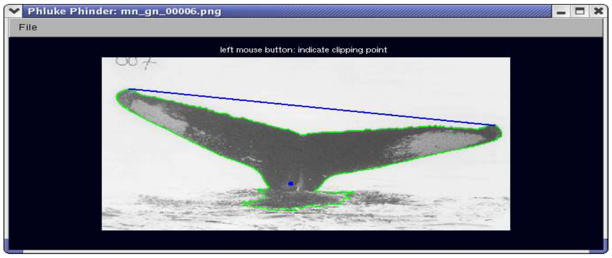
To extract the white and dark patches and the natural markings the user is prompted to specify three fluke *landmarks*, viz. the left and right fluke tip and the central fluke notch. These landmarks have also been used for photo-identification of flukes in [6].

Next, we use Otsu's grey-level thresholding [8] (which results in the optimal reduction of variance) on the extracted fluke to obtain an initial segmentation into dark and light patches. Finally, the interface supports local thresholding in order to allow the user to fine-tune this patch segmentation in regions of special interest where the global thresholding failed to catch subtle, but significant details (Fig.1 (c)) .



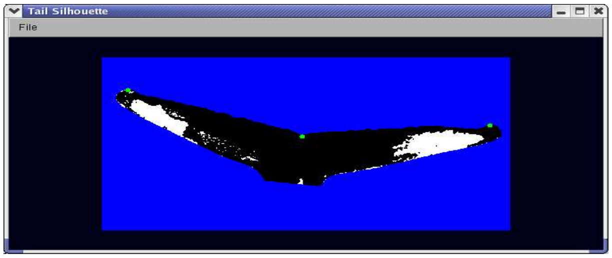
Left mouse button: add nodes - delete: remove last node - other buttons: finish polygon

(a) Original image and initial rough marker for tail.



left mouse button: indicate clipping point

(b) Watershed-based segmentation of tail (green contour). The tail is clipped at its base at a user-supplied point (blue) by fitting a line parallel to the blue line connecting the fluke tips.



(c) Final segmentation result -black and white patches used for the identification. The three landmark points are indicated in green.

Fig. 1. Tail and fluke patches extraction.

3. MATCHING

3.1. Fitting a Coordinate Grid

Images typically exhibit a large variation in viewing angles, distances and fluke inclination. In [5] it is argued that since fluke surfaces are nearly planar with dimensions significantly smaller than the distance to the camera, these variations can be modelled using affine transformations such as rotation, translation and scaling. It is important the developed methods are invariant up to a class of transformations since the invariants are an essential part of any single view reconstruction problem [4]. To be robust with respect to the above-mentioned variability, we therefore propose an affine-invariant method for computing a coordinate grid that can be superimposed and will divide the extracted fluke into N_R regions. The idea is very straightforward using a geometrical method for the grid construction. Similar ideas are employed in [4] for construction of an invariant patch identifier for a spotted dolphin identification.

A triangle LOR (Fig. 2) defined via the three preselected landmarks is constructed.

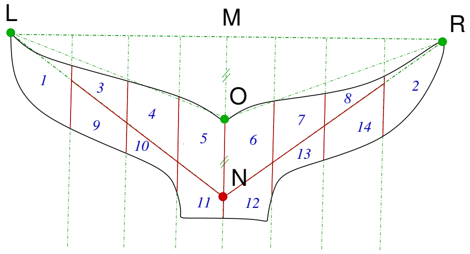


Fig. 2. Affine grid construction.

The base of the triangle (the line connecting the fluke's tips L and R) is divided in two equal parts by the point M. The symmetrical point of M in respect to O, i.e. N is found. $\triangle LOR$ is affine invariant and this is also true for $\triangle LNR$, which we use for the tail grid construction. Each fluke is then divided into n parts with lines parallel to the median NM. Thus, the grid delineates $N_R = 4n - 2$ (the tips are considered single regions) grid regions. These regions are labelled 1 through N_R starting from the fluke tips and then scanning left to right, top to bottom. For the grid on Fig. 2 $n = 4$, $N_R = 14$.

Notice that since the construction is solely based on affine invariant concepts (i.e. middle point, symmetry, equal distances, parallel lines), the resulting grid is invariant under affine transformations.

3.2. Feature extraction and comparison

After the grid has been fitted to the segmented fluke an N_R- dimensional feature vector $\mathbf{f} = (f_1, \ldots, f_{N_R})$ is computed. Each element f_i equals the ratio of the number of white pixels to the total number of pixels in the i-th grid region R_i. The feature vector for each fluke image is computed and stored in a database of features $\mathbf{F} = \{\mathbf{f}_1, \ldots, \mathbf{f}_N\}$ for all N images of the image database.

The matching process involves a comparison of the feature vector $\mathbf{q}$ calculated from a query image against all entries in $\mathbf{F}$. This is done by computing the average Euclidean distance per fluke segment

$$d(\mathbf{q}, \mathbf{f}) = \frac{\sqrt{\sum_{i=1}^{N_R} I_i (q_i - f_i)^2}}{\sum_{i=1}^{N_R} I_i}, \tag{1}$$

where the variable I_i indicates if the corresponding region of any of the pair of tails to compare should be considered, i.e.: $I_i = I_i^{\mathbf{q}} I_i^{\mathbf{f}}$.

The indicator for each fluke and its corresponding grid equals 1 for all regions above the clipping line c and 0 for the ones which are occluded. Because different regions will be occluded with the different tails we need to normalise the distance over the number of regions used for computation of the similarity of any pair of tails.

The images in the database are then ranked based on their similarity to the query image.

4. RESULTS

A database of 69 unique grey-scale images of humpback whale flukes was available for testing the proposed methodology. The images were of different resolution and different quality which is common in marine biologist practise. The database has been manually processed by an expert and 32 individuals were identified. For 5 of these individuals there were 3 different images available (triple) and for the rest (27) there were 2 images (pair) for each individual.

4.1. Fluke and patches extraction

The watershed segmentation provided an excellent contour of the tail for most of the data in one iteration. For the remaining images (mainly of poor quality), the user could achieve very good extraction after few iterations of fine-tuning by interactively introducing additional markers using the GUI. The semi-automatic watershed approach showed overall a very good performance and was able to extract the flukes for all images. The subsequent thresholding (automatic or interactively fine-tuned) produces a very good binary representation of the flukes and the natural markings.

Figure 3 illustrates the performance of the flukes patches extraction for several images of the database. It can be seen that the segmentation captured the important markings irrespective of the original image quality or level of difficulty.

4.2. Grid fitting and Matching

Figure 4 illustrates grids ($n = 8, N_R = 30$) fitted to the segmented pairs of some images, which were identified by an expert as matches. It should be noted that salient markings appear in the correct grid region independently of the viewing angle and tail slant, especially visible for the second pair of tails.

Two matching strategies were tested. The first one ignores all completely or partially (i.e. $I_i = 0, \forall i : R_i \cap c \neq \emptyset$) occluded regions. The second strategy ignores only the occluded regions outside $\triangle LNR$. The second method achieves better retrieval results on average as summarised in Table 1.

Table 1. Percentage of individuals whose true match is ranked among the top k.

$I_i = 0, R_i \cap c \neq \emptyset$	$k = 10$	$k = 3$	$k = 1$
$\forall i$	94.2	78.2	60.8
$i, R_i \notin \triangle LNR$	100	82.6	69.5

Both strategies reduced the number of images which had to be reviewed by the expert by a factor of 7. All images had their true match ranked amongst the top 10 using the second strategy and for nearly 70% of the database the true match(es) were correctly identified as the first (or first and second in case of triples). Some more difficult cases are illustrated in Fig. 5 where the true match of a query image belonging to a pair was ranked third and in Fig. 6 where the true matches of a query from a triple were ranked first and third. It can be noted that the "false positives" are still visually similar to the query image.

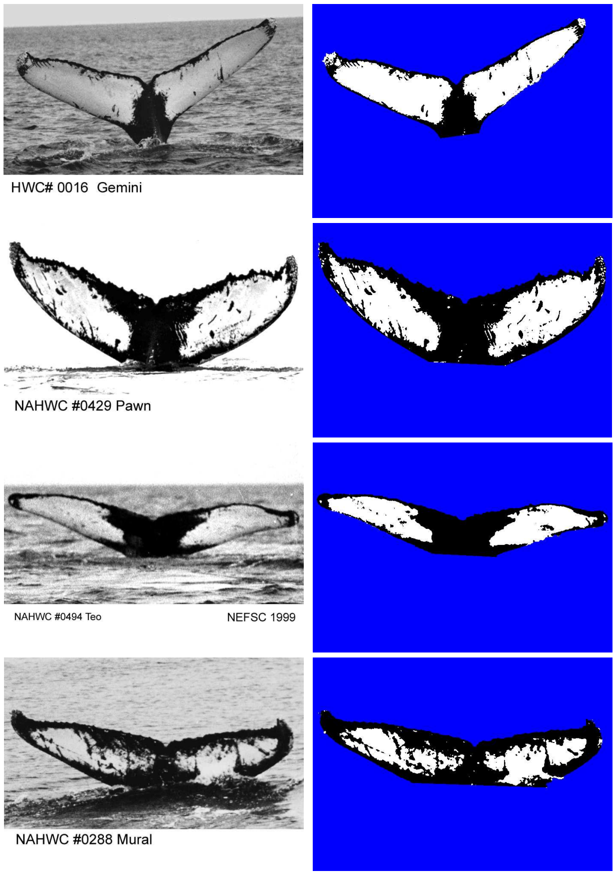


Fig. 3. Fluke and patches segmentation. Original image (left) and its segmentation (right).

5. DISCUSSION

The work reported in this paper addresses two aspects of the photo-identification problem: the *segmentation* of relevant image information (fluke, patches) and the *feature extraction and matching* based on an affine invariant coordinate grid.

The *segmentation* program has been tested by marine biologists during a Europhlukes Project software evaluation test meeting, where it was received favourably.

The performance of the *grid-based feature extraction and matching* needs to be confirmed on a much larger database, and we are currently setting up further test protocols. However, it bears pointing out that the methodology is quite generic and can easily be extended to other photo-identification problems where similar markings are used to distinguish between individuals (e.g. dorsal fins for dolphins).

Results suggest that a more elaborate feature vector construction should be a main direction for further research efforts. For instance we are currently working on a salient pattern detector.

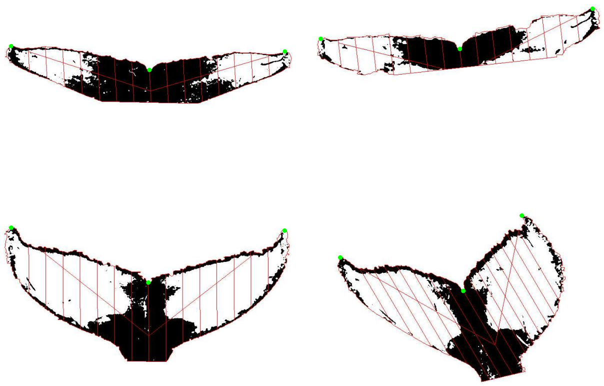


Fig. 4. Grid fitting to pairs of images (a pair per row), which are known to be of the same individual. The landmarks are shown in green.

The descriptors of the salient markings in respect to the affine coordinate system, could dramatically improve the recall specificity.

ACKNOWLEDGEMENTS: This work has been partially supported by Europhlukes (www.europhlukes.net, EC project-ID EVR1-CT2001-20007). J. Allen from College of the Atlantic, USA is gratefully acknowledged for providing the images.

6. REFERENCES

[1] B.N.Araabi, "Syntactic/Semantic Curve-Matching and Photo-Identification of Dolphins and Whales", PhD Thesis, Texas A&M University, 2001.

[2] B.N.Araabi et al., "Generalization of Dorsal Fin Ratio for Dolphin Photo-Identification", *Annals of Biomedical Engineering*, vol. 28, no. 10, Oct. 2000, pp. 1269-1279.

[3] P. Soille, "Morphological Image Analysis", *Springer*, 2003.

[4] B.N.Araabi at al., "Locating an Affine/Projective Invariant Identifier Patch on an Image", *Fifth IEEE Southwest Symposium on Image Analysis and Interpretation*, pp.121-125, April 2002.

[5] N. Kehtarnavaz et al., "Photo-identification of Humpback and Gray Whales Using Affine Moment Invariants", *Proceedings of 13th Scandinavian Conference on Image Analysis*, Sweden, July 2003.

[6] G. Hillman et al., "Computer-Assisted Photo-Identification of Flukes Using Blotch and Scar Patters ", *Proceedings of 15th Biennial Conference on the Biology of Marine Mammals*, US, Dec. 2003.

[7] S. Mizroch et al., "Computer Assisted Photo-Identification of Humpback Whales", *Individual Recognition of Cetaceans*, International Whaling Commission, Cambridge, 1990.

[8] N.Otsu, "A Threshold Selection Method from Gray-Level Histogram", *IEEE Transactions on System, Man, Cybernetics*, vol. SMC-9, 1997, pp.62-66.

[9] WhaleNet http:/whale.wheelock.edu

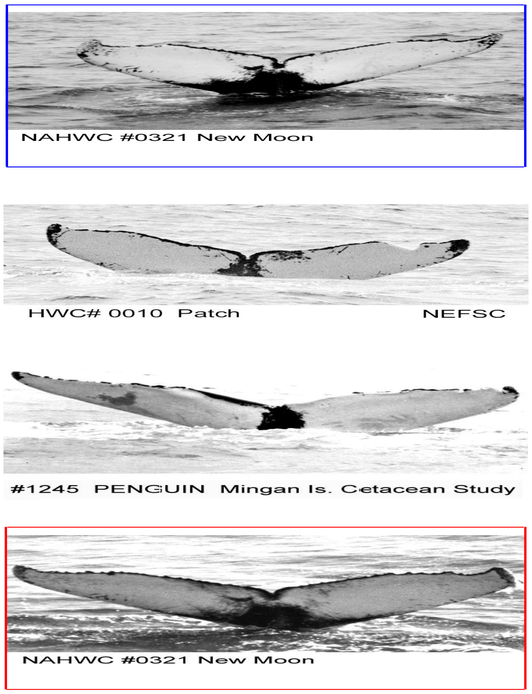


Fig. 5. Query image (top, framed in blue) and first three matches. The query is part of a pair and its true match is framed in red.

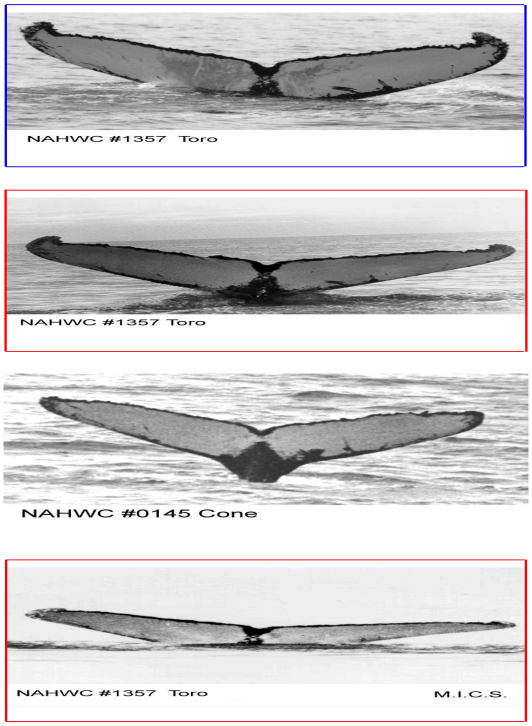


Fig. 6. Query image (top, framed in blue) and first three matches. The query is part of a triple and its true matches are framed in red.