# Iterative 3-D Pose Correction and Content-Based Image Retrieval for Dorsal Fin Recognition⋆

John Stewman, Kelly Debure, Scott Hale, and Adam Russell

Eckerd College, St. Petersburg FL 33712, USA

**Abstract.** Contour or boundary descriptors may be used in content-based image retrieval to effectively identify appropriate images when image content consists primarily of a single object of interest. The registration of object contours for the purposes of comparison is complicated when the objects of interest are characterized by open contours and when reliable feature points for contour alignment are absent. We present an application that employs an iterative approach to the alignment of open contours for the purposes of image retrieval and demonstrate its success in identifying individual bottlenose dolphins from the profiles of their dorsal fins.

## 1 Introduction

Marine mammalogists study numerous characteristics of their target populations, estimate population size and longevity, identify range limitations and make observations on behaviors and association patterns. Identification of individual animals is essential to these studies. For decades the identification method of choice has been photo-identification. This technique is far less invasive than traditional mark and recapture methods which used tagging or freeze branding, and a wealth of additional information is captured in the identification photographs acquired during field studies.

Bottlenose dolphins are very active, social, sometimes aggressive animals. Their dorsal fins acquire damage over time, particularly in the form of nicks, notches, scratches, and other markings. As Figure 1 shows, these features are sufficient to identify individual adult dolphins. This is similar to the use of a biometric identifier, such as a fingerprint, to recognize a person, as an alternative to token-based identification requiring ID cards or PIN numbers.

This paper focuses on the application of image processing and analysis techniques to semi-automate the process of photo-identification. The methods described should apply equally to other species of marine mammals that have dorsal fins of significant extent. The application described here has been specifically developed and tested using images of the dorsal fins of bottlenose dolphins (*Tursiops truncatus*).

**Fig. 1.** Dolphin Dorsal Fin with Distinctive Natural Markings

## 2   Background

Researchers photograph dolphins in their natural surroundings and compare new photographs from each sighting against a catalog of reference images of known dolphins in order to verify the identity of the recently sighted individual. The catalogs, in either photo, slide or digital form, may be organized into categories of distinct fin shapes and each of the shape categories may be further divided into subcategories denoting regions where the predominant fin damage is present. The manual photo-identification process, although effective, is extremely time consuming and visually stressful, particularly when large collections of known dolphins are involved. In some studies, the tentative identification, or the finding of a truly new dolphin not in the catalog, must be confirmed by a second researcher, replicating the same image search. The sheer quantity of image data and the somewhat constrained properties of this class of images makes the automation of the photo-identification process an appealing and very feasible subject for content-based image retrieval. The complications associated with such a project include the difficulty of creating a reliable and intuitive method for database query and the inherent computational intensity associated with the registration and comparison of image data.

Methods for the computer assisted identification of individual marine mammals have been proposed for a variety of different species including gray seals[5], sperm whales[15], [7], dolphins[4], [10], [1], humpback whales[12], [14] and right whales[2], [6]. Identifying dolphins by the comparison of dorsal fins is similar to whale identification based upon fluke characteristics such as highly visible pigmentation patterns. While more subtle pigmentation and injury patterns such as rake marks may be used as secondary identifiers, it is the overall shape and the specific details of the damage found along the fin profile that have become the primary identifiers used within the cetacean research community to identify individual bottlenose dolphins. Although methods have been developed which use the profile of the trailing edge of whale flukes for individual indentification, characterizing the outline of a dorsal fin has some unique challenges. Dolphin dorsal fin outlines have no easily determined beginning and ending points. The user must arbitrarily decide where the dorsal fin and the dolphin's back meet. Often the dorsal fin is partially obscured by waves or water if the dolphin is diving or surfacing. This uncertainty leads to significant variation in the portions

of the fin outline that are used to determine matches and creates complications in scaling and alignment of outlines.

The Digital Analysis and Recognition of Whale Images on a Network (DAR-WIN) system is in many aspects, a straightforward application of image processing, computer vision, and content-based image retrieval. The novelty of the DARWIN system, in its current form, is in the use of well known techniques to provide an almost completely automatic process, from image load to display of a rank ordered list of possible dolphin identities, while also providing numerous user check points and correction tools, allowing the user to enhance or even override the system's automatic processes. This produces significant time savings, and still leaves the human user in the position of final decision maker.
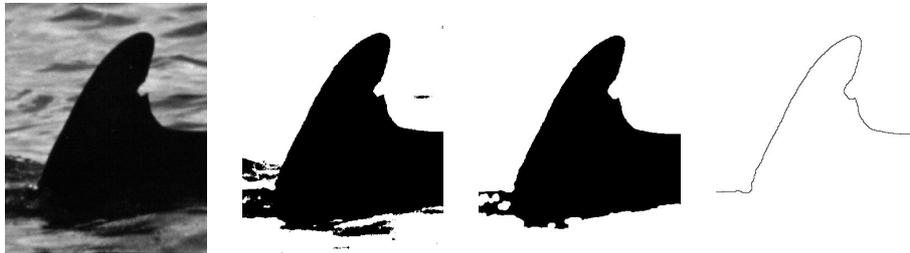


**Fig. 2.** Left: original image; Center left: image which results from unsupervised selection of threshold value; Center right: smoothing of the fin region and removal of small regions with morphological processing; Right: the resulting outline

## 3    Sketch-Based Image Query

A researcher wishing to identify an unknown dolphin must open a digital image of a dolphin dorsal fin and orient the image such that the leading edge of the dorsal fin is at the left of the image. The software automatically generates an outline of the dorsal fin. It uses an unsupervised minimum error threshold selection [9] to select an intensity threshold that will effectively segment the image into dorsal fin and background regions. It then uses morphological processing to simplify the region of interest and generate its outline, as shown in Fig. 2. If the software is unable to identify the dorsal fin region in the image, the user has an opportunity to roughly trace a general outline of the leading and trailing edges of the dorsal fin with the cursor. The trace of the outline, automatically generated or hand-sketched, initializes the positions of a series of evenly spaced points along the edge of the fin and provides a constraint space for their automated repositioning. DARWIN uses active contours [8] to move the points from their initial locations to the actual edge of the fin. The movement is controlled by several energy measures. Large magnitude values in the image gradient attract the points, while internal energy measures that maintain even point spacing and curve smoothness restrict point movement. Figure 3 illustrates the movement
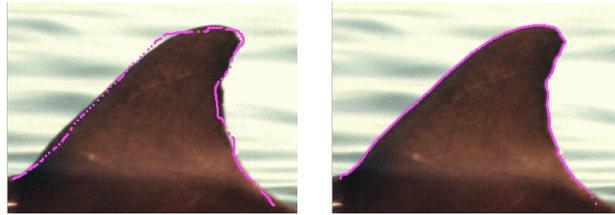
**Fig. 3.** Left: Initial location of outline; Right: repositioned outline using active contours

of the chain of points. Once the points have been appropriately repositioned, the user has the option to manually relocate individual points along the edge of the fin. This opportunity for user interaction is necessary since photographic artifacts such as glare spots are interpreted by the program as nicks and notches and poor contrast or intensity variation on the surface of the dorsal fin may necessitate manual positioning of the outline. In addition, parameters of the active contour which encourage the convergence toward a smooth outline, may cause the automatic repositioning to miss nicks or notches with extreme angles.

Several processes, including active contours, feature point location, and mapping of unknown to reference outline, have configurable constants that are set to values based on an assumption of approximate overall fin size. The actual values of these constants and the standard fin size have been determined empirically, such that the best results are obtained in all processing of the fin outlines. All fin outlines are scaled to an approximate height of 600 units. The height measure is computed and scaling is applied before the locations of feature points are known. Therefore, the two points used to compute the height give only an approximation of the vertical distance between the dolphin's back and the tip of the fin. As Figure 4 shows, $Q$ is the midpoint of a line segment connecting the first outline point and the last outline point, and $R$ is the outline point with the index $\frac{n}{2}$, where $n$ is the total number of points along the outline. $Q$ is a reasonable estimate of the midpoint along the dolphin's back where the fin is attached, and $R$ is a point in the vicinity of the tip of the fin. A scale factor
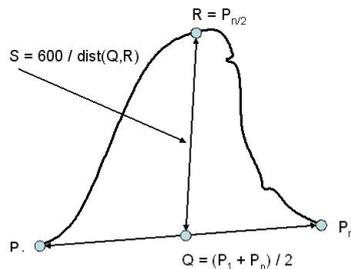


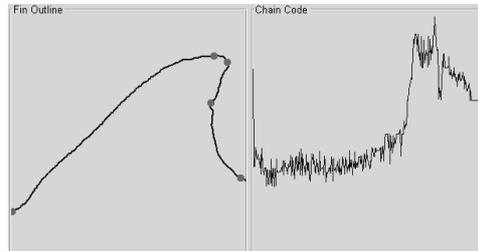**Fig. 4.** Scale factor computation for the standardization of outline size

**Fig. 5.** Left: Dorsal fin outline with feature points; Right: Chain code of outline

$s = \frac{600.0}{dist(Q,R)}$ is multiplied times the image coordinates of each outline point in order to standardize the outline size. After the outline size is standardized, the outline (interpreted as a polyline) is traversed and a new sequence of outline points is computed with an approximate, uniform point to point separation of 3.0 units. This typically produces an outline containing 400 to 600 points. The reason the point spacing is approximate, is that the even spacing computations are done while the outline is still in an integer coordinate form, so all point coordinates are rounded to the nearest integer value.

## 4    Feature Point Identification

The DARWIN system represents each fin with an outline contour (a sequence of evenly spaced points approximating the its 2-dimensional placement and shape), a chain code (a one dimensional representation of the orientation of successive outline edge segments), and a set of salient features (Tip, Notch, ... ). These feature points are used to establish intial alignment of two fin outlines during matching, and they are identified by an unsupervised process using the chain codes extracted from the contour. A single process converts the initial outline contour to even point spacing and creates the chain code.

Dorsal fin outlines contain only a few consistantly identifiable features suitable for alignment. As illustrated in Figure 5, the DARWIN system automatically locates 1) the start of the leading edge, 2) the end of the leading edge (where the leading edge angle begins to bend toward the fin tip), 3) the tip of the dorsal fin, 4) the most prominent notch on the trailing edge, and 5) the end of the trailing edge.

Figure 5 shows the general shape of the plot of a fin's absolute chain code as that of a step edge and the location of the tip corresponds to the point of transition from low to high. DARWIN uses a quadratic spline wavelet [11] which produces large coefficients for step edge type features in the transformed signal. In order to identify the tip of the dorsal fin, a wavelet decomposition of the chain code comprising the outline is computed. At the coarsest level of analysis, the tip is identified by locating the maximum coefficient value. This initial positon is tracked back through each of the finer levels of the decomposition, to more
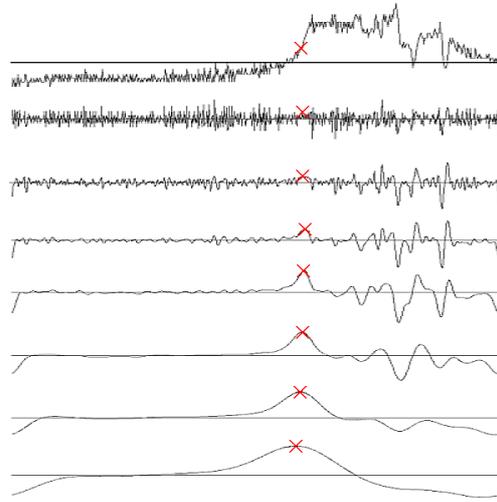
**Fig. 6.** Finding the dorsal fin tip from the wavelet transform of the outline. A plot of the original chain code is shown on top, with increasingly coarse details below. The position of the tip is found on the coarsest (bottom) level, and tracked back through the finer levels to obtain a more accurate location.

precisely identify the position of the tip in the chain code [11], as illustrated in Fig. 6.

Notches along the trailing edge of a dolphin dorsal fin are often the primary features used in the manual photo-identification process. Prominent notches in a fin outline appear as valleys in the plot of a fin's absolute chain code and a wavelet decomposition will produce local minima for features of this type. Since it is the most prominent notch along the trailing edge of the dorsal fin that is desired, only that portion of the outline following the fin tip is analyzed. Local minima with the largest magnitudes are identified at an intermediate transform level as candidate notches and are tracked to coarser levels. The minima that decreases most slowly in magnitude across the coarser levels is selected as the most prominent notch. As with the fin tip, the precise location of the notch is determined by back-tracking through increasingly finer levels of the decomposed signal.

To automatically identify the starting point of the dorsal fin outline at the base of the leading edge, the absolute angles between successive points along the outline are examined. A threshold angle is selected to maximize between class variance of the angles which comprise the proper leading edge of the fin and the angles which are associated with the body of the dolphin. This approach is based on Otsu's method for unsupervised threshold selection in image data [13]. Line segments at the leftmost end of the outline are discarded if their angles diverge significantly from the predominant orientation of the leading edge, accurately identifying the point at which the dorsal fin meets the dolphin's back. If this junction is occluded or missing from the initial outline, the first point of the

edge is stored as the start of the dorsal fin, indicating a truncated trace. The ending point of the leading edge is similarly identified as the point at which the angles of the line segments begin to significatly diverge from the predominant orientation of the leading edge.

## 5   Mapping Unknown Fin Outline to Reference Fin Outline

When photographing dolphins in the wild, there is little control over the orientation (3-D pose) of the fin that appears in an image. Several constraints are imposed, and must be assured by the user. The fin must be upright, and the dolphin must be swimming to the left. Tools are provided in the software to allow the user to flip the image as needed to assure this. Since the software is only concerned with the outline or profile of the fin, it does not matter that an image is inverted and processed as if an image of the right side of a dolphin's dorsal fin were an image of its left. These modest constraints simplify the process used for feature point location, ensuring that the trailing edge of one dolphin is compared to the trailing edge of the other. In order to correct for the differences in 3-D pose between two fin outlines ($F_i$ and $F_j$), one is mapped to the other using a three point affine transformation. The three points used are selected from the previously described identifiable fin features and are guaranteed to be non-colinear. The 2-D coordinates of the six feature points, taken from the scaled, evenly-spaced fin outlines define two triangles ($T_i$ and $T_j$) in the image coordinate space. A transformation matrix is found that maps $T_i$ to $T_j$. This defines a skew transformation in the 2-D image plane, but it also corresponds to many 3-D affine transformations that map $T_i$ to a 3-D pose that projects to $T_j$. Given several reasonable assumptions, this three point transformation gives a close enough approximation in the image plane, to a complete 3-D pose correction between the two fin outlines. These assumptions are :

 – Actual 3-D pose is unimportant, only the corrected projection in the image plane is needed.
 – The three feature points are reliably detectible
 – Foreshortening effects are negligible
 – The more similar two fin outlines are, the more accurate will be the pose correction/estimation

Once the affine transformation matrix is found, it is used to map all points on the unknown fin outline $F_i$ to the orientation of the known fin outline $F_j$.

The degree to which two fin outlines, or their open polyline approximations, match is measured by a mean squared error computed using "corresponding points" along some fixed extent of the two outlines. First, the error measure is used in an iterative process to refine the alignments of the two outlines such that a better fit is obtained. The better fit is defined simply as the fit that produces the smaller mean squared error. Second, the ranking order of final match results is established by using the mean squared error for the best alignment of each

pair of reference and unknown outlines. The listing is in order of increasing mean squared error. The mean squared error measure has been satisfactory for purposes of alignment, but has been less than optimal for determination of the final rankings. Future enhancements to the software will employ additional measures of fit that allow more local, small scale analysis of particularly salient areas of the outlines. The use of more localized error measures would more closely emulate the discriminations made intuitively by biologists in their comparisons of two dorsal fins.

The original three point mapping was done once per fin pair, using the beginning of the leading edge, the tip and the position of the most significant notch as the alignment points. This produced reasonable results as long as the entire leading edge of the fin was visible in both images and was fully part of the outline, as long as the dorsal fin tip was well defined (not true if severely nicked or missing), and as long as the most significant notch was well defined and positioned at a significant distance from the tip. Full fin outlines from two images of the same dolphin, extending between correspondingly placed beginning and ending points and having a clear most significant notch, always produced excellent quality mappings. However, misplacement of the beginning of the leading edge by even relatively small amounts (10% of leading edge length) produced significant misalignments. Also, a significant notch close to the tip produced a very skinny triangle, such that only a few pixels of misplacement caused significant misalignment. Finally, it was found that the multiscale approach to notch detection was sensitive to the overall length and curvature of the trailing edge. Trailing edge traces that include part of the back could bias notch detection creating a preference for notches closer to the end of the trailing edge. The undesirable effects of highly variable user input indicated a need to move away from single applications of a pose correction mapping based upon these initial feature points.

The first significant improvement to the alignment of fin outlines and rankings of matches was through a "best of 13" approach. In this approach, the original mapping based on the feature points is applied, and the mean squared error is computed between the mapped unknown outline and the known outline. Then twelve alternative mappings are performed and error measures computed. Six of the alternatives test for shortened traces of the unknown fin outline and represent shortening of the leading edge in increments of 5%. Similarly, the other six alternatives test for shortened traces of the known fin outline. The mapping that produces the smallest mean squared error is selected to correct for orientation differences between the two fin outlines. In practice, this approach produced significantly better mapping results. Overall performance is measured by how close the correct "known" dolphin is to the top of the ranked listing of the database when queried with an "unknown." The median position of the correct identity using the "best of 13" approach was 8/200 or 4% down from the top of the list. The mean location of the correct identity was 22/200 or 11% down from the top. However, sensitivity to misplacement of the feature points at the notch and at the end of the trailing edge still produced numerous intuitively poor

mappings. The desire has always been to produce correct rankings of candidate identities, and to produce intuitively reasonable orientation corrections, even when the fin being matched is not that of the same dolphin.

Several iterative optimization approaches have been implemented in the software and tested. These all share a basic Newton-Raphson approach. Three point correspondences are used to map the unknown fin outline to the known fin outline in all cases. Shortening of leading and trailing edges, and movement of the tip, have all been done by testing feature point placement corrections of 1% of the index range of the fin leading edges and then making jumps of up to 16% based on the best correction indicated during each iteration. The maximum feature movement amount begins at 16% and decreases by half under various conditions, until either none of the 1% tests indicate improvement or the jump interval decreases to 0%. The major difference between all of the variations developed and tested is in the choice of which points to move.

Our current optimized approach allows the feature points at the begining of the leading edge and the end of the trailing edge to move, shortening the outline. This repositioning is performed on both the unknown and reference outlines, and continues as such until the jump interval decreases to 4%. At this point, the feature point for the tip of the dorsal fin of the unknown dolphin may be shifted as well, to further reduce mean squared error between the outlines.

Since the jump intervals begin at 16% of the index range of the fin leading edges, and they monotonically decrease when a direction of improvement is found but the jump in that direction is too far to improve (decrease) the error, the optimization process is guaranteed to terminate. It is possible that the process will settle into a local minimum of the error function. This has been observed in practice. However, the use of the auto-trace feature to find the initial placement of the fin outlines appears to minimize occurrences of non-optimal minima being found.

## 6    Selection of Corresponding Points for Error Calculation

When comparing open curves ($C_1$ and $C_2$), or even their polyline approximations, it is difficult to calculate the location of the point on $C_2$ that best corresponds to a given point on $C_1$. In general, the best corresponding point is the point on $C_2$ that is closest to the current point on $C_1$, but located in the same direction of travel along $C_2$ as is the given point on $C_1$ from the previously selected pair of corresponding points. Essentially, this means that the entire sequence of points traversed along $C_1$ moving in the same direction has a single best set of similarly ordered corresponding points on $C_2$. For the purpose of calculating a mean squared error between two mapped outlines, it is not essential that one solves exactly for the best corresponding points along the pair of curves. When alignments are near exact, error based on a reasonable point-to-point correspondence should be zero or nearly so. When alignments are not close, the outline curves are most likely from fins of different dolphins. In these cases it is acceptable to have point correspondences that are not exactly ordered in sequence on
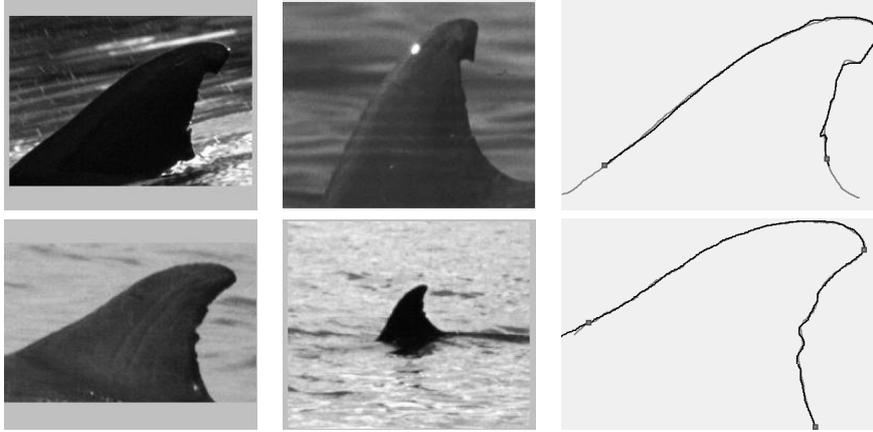
**Fig. 7.** For each set of images above, Left: the reference image from the database, Center: the unknown fin from a sighting, and Right: the registered outlines, unknown mapped to database

both curves. Such situations can be expected to produce larger errors in many cases, but this is quite acceptable since the outlines of different dolphins should not be given a minimal mean squared error match in any case. If the error calculation is slightly less than actual, this is acceptable as well, since the overall error magnitude of such grossly misaligned fin outlines will produce a large error and be ranked low in the final listing regardless. The point here is that an acceptable selection method should produce a sequence of corresponding points on the two curves that is "good enough" to produce a mean squared error suitable for use in fin outline registration and fin match ranking. The currently used method does just that.

It is important to note that an outline curve which is evenly spaced and similarly scaled to another outline before mapping (orientation correction) is no longer evenly spaced after mapping. Thus, the outlines cannot simply be traversed by using similar curve point indices to select individual points for comparison. Mapping is based on a 2-D skew transform that only solves for the approximate 2-D projection of a 3-D affine transformation of the mapped curve. Therefore, the sequence of line segments connecting similarly indexed points on the two curves, even after trimming extra points on curve ends, has the look of the ribs of an accordion and the calculated error, even when curves are well aligned, can be significantly higher than desired.

The current method of selecting corresponding point pairs uses an arc length based approach. The total arc length of each curve is computed between its repositioned endpoints. The ratio of the two arc lengths is used to determine the distance to traverse on the unknown curve for each unit step on the known curve. The known curve is then traversed using the indexed curve points. At

each point P(i) at arc length L(i) from the beginning of the curve, a point Q(i) is found on the mapped unknown curve at L(i)*ratio distance from the beginning of the unknown curve. The midpoint of the segment connecting P(i) and Q(i) is a point M(i) on a medial axis of the region in the image plane bounded by the two outline curves. This sequence of medial axis points, M(i) for i = 0..m, is then traversed. At each medial axis point, a perpendicular to the medial axis is intersected with the two fin outline curves. The two points of intersection are the $i^{th}$ pair of "corresponding" curve points used in the error computation. The squared length of each segment between the m corresponding point pairs forms the basis of the mean squared error. In practice this approach yields closely corresponding point pairs. They are not the optimal pairs in all cases, but are "close enough" to produce meaningful error measures. Alignments based on these pairs are tight and intuitive as shown in Fig. 7.

## 7    Testing and Results

Preliminary testing of the software utilized a database of dorsal fin images for 200 individuals. Fifty additional images of individuals known to be contained in the database were used as "unknown" fins to test the query mechanism. Both the database images and the fifty previously identified dorsal fin images represented a range of quality and resolution and depicted fins with damage of varying location and degree. As shown in Figure 8, queries performed with forty of the fifty "unknown" fins produced rankings of the corresponding correct fin in the top 10% of the database. Of these, twenty three were ranked in the first position. On average, the optimized alignment method placed the correct fin in the top 9% of the ranked listing of database fins. The median ranking was in the top 1%. The outlines of the "unknown" fins were well aligned with the outlines of their counterparts in the database, thus the registration method
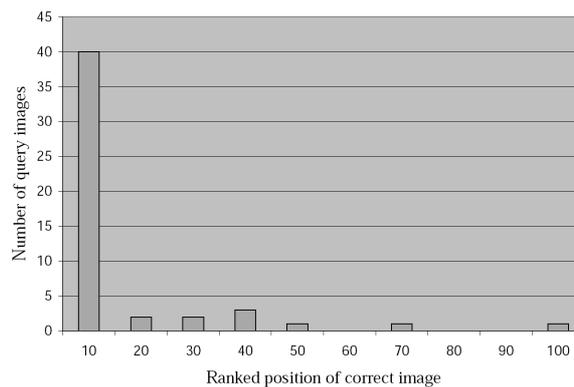


**Fig. 8.** Rankings of the correct fin in the database as a percentage of the total number of database entries

produced accurate alignments, even for images photographed from significantly different camera angles, as shown in Fig. 7. Also, in most cases, the outlines of the "unknown" fins were well aligned with outlines of dorsal fins of images of different individuals. Intuitive alignment is important even for the non-matched cases to give biologists the opportunity to visually compare pairs of dorsal fins removing the distortions that arrise from difference in distance and camera angle. Of the test set of 50 "unknown" fins, only one of the corresponding correct fins from the database was ranked worse than 100 (random selection).

These results suggest that the registration methods presented herein show significant promise in perspective correction of dolphin dorsal fin images. The iterative nature of the alignment algorithm makes it less sensitive to the original positioning of the outline or subsequent identification of feature points. In any case, the process has the ability to improve the registration of outlines considerably. The improved registration makes fin outline comparisons easier and aids in the subsequent retrieval of appropriate images.

## References

[1] B. N.Araabi, N. Kehtarnavaz, G. Hillman, and B. Wursig, "A string matching computer-assisted system for dolphin photo-identification," *Annals of Biomedical Engineering* **28**, pp. 1269–1279, 2000.

[2] S. R. Burnell and D. Shanahan, "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 Supplement* **2**, pp. 297–300, 2001.

[3] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **8**, pp. 679–698, 1986.

[4] R. H. Defran, G. M. Shultz, and D. W. Weller, "Technique for the photographic identification and cataloging of dorsal fins of the bottlenose dolphin (tursiops truncatus)," *Report to the International Whaling Commission* **Special Issue 12**, pp. 53–55, 1990.

[5] A. R. Hiby and P. Lovell, "Computer-aided matching of natural markings: A prototype system for grey seals," *Report to the International Whaling Commission* **Special Issue 12**, pp. 57–61, 1990.

[6] A. R. Hiby and P. Lovell, "A note on an automated system for matching of callosity patterns on aerial photographs of southern right whales," *Journal of Cetacean Research and Management Supplement* **2**, pp. 291–295, 2001.

[7] R. Huele and H. U. de Haes, "Identification of individual sperm whale flukes by wavelet transform of the trailing edge of the flukes," *Marine Mammal Science* **14**, pp. 143–145, January 1998.

[8] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision* , pp. 321–331, 1988.

[9] J. Kittler and J. Illingworth, "Minimum Error Thresholding," *Pattern Recognition*, Vol. 19, No. 1, pp. 41-47, 1986.

[10] A. Kreho, N. Kehtarnavaz, B. N. Araabi, G. Hillman, B. Wursig and D. Weller, "Assisting manual dolphin identification by computer extraction of dorsal ratio," *Annals of Biomedical Engineering 27*, pp. 830-838, 1999.

[11] S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Transactions on Pattern Analysis and Machine Itelligence*, Vol. 14, No. 7, pp. 710-732, July 1992.

[12] S. A. Mizroch, J. A. Beard, and M. Lynde, "Computer-assisted photo-identification of humpback whales," *Report to the International Whaling Commission* **Special Issue 12**, pp. 63–70, 1990.

[13] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-9, No. 1, pp.62-66, 1979.

[14] R. K. Uyeyama and L. M. Herman, "Automated fluke-contour matching of humpback whales using a wavelet-based algorithm," 14th Biennial Conference on the Biology of Marine Mammals *(abstract)*, Vancouver, p. 221, 2001.

[15] H. Whitehead, "Computer-assisted individual identification of sperm whale flukes," *Report to the International Whaling Commission* **Special Issue 12**, pp. 71–77, 1990.