Marine Mammal Science

MARINE MAMMAL SCIENCE, 26(3): 744–756 (July 2010)

DOI: 10.1111/j.1748-7692.2009.00368.x

Fluke Matcher: A computer-aided matching system for humpback whale (*Megaptera novaeangliae*) flukes

ERIC KNIEST
School of Engineering,
University of Newcastle,
Callaghan, NSW 2308, Australia
E-mail: eric.kniest@newcastle.edu.au

DANIEL BURNS
PETER HARRISON
Southern Cross University Whale Research Centre,
School of Environmental Science & Management,
Southern Cross University,
Lismore, NSW 2480, Australia

Humpback whales (*Megaptera novaeangliae*) are characterized by variable natural pigmentation patterns and scarring marks on their tail flukes and other regions of the body (Lillie 1915, Schevill and Backus 1960, Katona *et al*. 1979, Chu and Nieukirk 1988). The ability to recognize individual humpback whales from photographs of their tail fluke pigmentation and scarring patterns was first realized by researchers in the 1970s (Katona *et al*. 1979). Since that time, the technique of photo-identification has been widely used on humpback whale populations around the world to determine many aspects of their biology, ecology, and behavior.

However, a significant and growing problem exists with comparing photo-identification data sets. Traditional methods of matching photographs of ventral fluke surfaces require manual pairwise comparison of all images within and among data sets, which are often very large. This process requires substantial time, effort, and expertise. Furthermore, as each of the data sets grows, the number of comparisons required increases exponentially.

Computer-based matching systems have been developed for other marine species, including some cetaceans (Araabi *et al.* 2000, Huele *et al*. 2000, Hiby and Lovell 2001, Arzoumanian *et al.* 2005, Beekmans *et al.* 2005, Caiafa *et al.* 2005, Speed *et al.* 2007, Van Tienhoven *et al.* 2007). However, attempts to develop a computerized system for humpback whales are still in their infancy and have focused largely on patch distribution of black and white markings on the fluke (Kehtarnavaz *et al*. 2003, Ranguelova *et al*. 2004). To date, none of the systems used for computer-aided fluke identification employ the wide range of features and unique properties of humpback whale flukes that can be utilized to more efficiently identify individuals.

Substantial research on humpback whales over the past few decades has resulted in the collection of many thousands of images of humpback whales for photo-identification purposes. Here, we present a powerful, new technology for analyzing

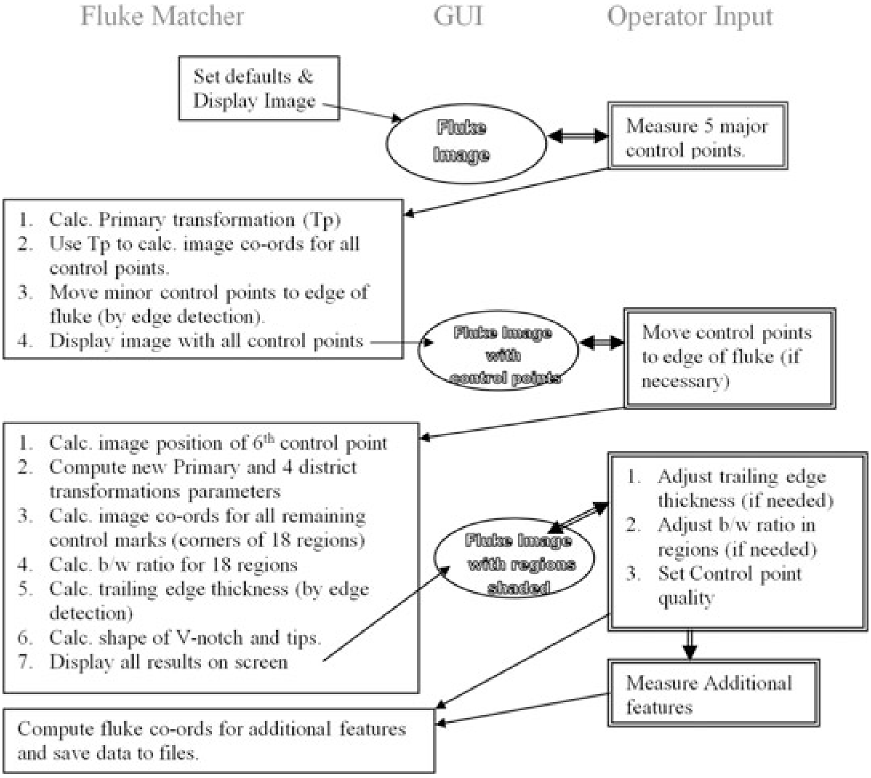

Figure 1. Flow chart for measurement (input) phase of Fluke Matcher. Left side is program processing, middle is Graphical User Interface (GUI), right side represents operator input.

photo-identification data that will greatly enhance the ability to reconcile large fluke catalogs. The system uses a wide range of criteria, based on multiple key features of humpback whale flukes that are normally utilized by manual matching methods, to produce a reliable matching system. These include the measurement of key features of the fluke, including parameters to describe the shape of the fluke, black and white pigment distribution in different regions of the fluke, and other distinctive features that enable identification.

One aim of the system is to utilize most of the photographs in existing catalogs. The quality of photographs in these catalogs varies greatly, with a large percentage of the photographs within any catalog having images of flukes that have large rotations or have a proportion of the tail underwater. Thus, the Fluke Matcher software does not rely on only one aspect of the fluke, such as the shape of the trailing edge, as this aspect is not visible with sufficient detail on many photographs.

The Fluke Matcher system consists of two main modules. The measurement module (or input phase) provides a user-friendly graphical interface where the operator can input new fluke photographs and then save it to the database (Fig. 1). The search module uses these data to find any matches in the existing fluke database. Fluke Matcher has been written in FORTRAN on a Windows-based PC and has

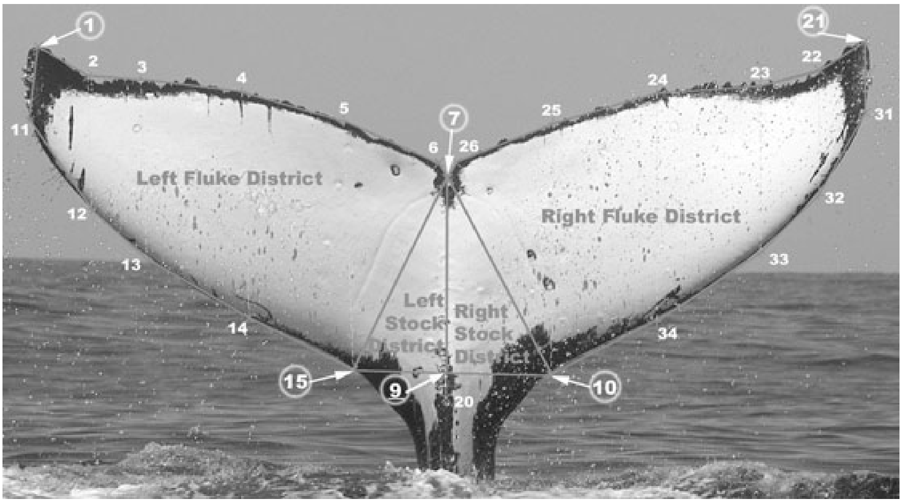

Figure 2. Diagram showing the four transformation districts and the positions of major (circled) and minor control points. Additional points (8, 16, *etc.* not shown) are used to define the 18 b/w ratio regions.

approximately 10,000 lines of code. It has also been run successfully on Macintosh computers using the Windows interface "Parallels." The program performs better with the use of a high-resolution monitor to enable closer inspection of each fluke photograph.

The first step for any computer-aided matching system is to measure or extract the data from each photograph. As part of this process, it is important to transform the imaged fluke onto a common reference system. This is accomplished in Fluke Matcher by the operator marking the position of five major control points on the photograph (Fig. 2). These points include the most easily identified points, such as the fluke tips and central V-notch, as well as two control points at the leading edge of the fluke. Because only a single photograph is used, the true size and shape of the fluke cannot be determined without the benefit of stereo-photography. Thus, each fluke measured is rescaled to one standard size (5 m from tip-to-tip to coincide with Tomilin (1957), who described the fluke width of humpback whales as approximately one-third of the total body length). This also provides a basis for comparing the same fluke over time, since its actual size may have increased. The transformation from the image co-ordinates onto the reference system can be done using a number of techniques that translate, rotate, and scale the image. To be able to perform a transformation of a single photograph, the fluke of a humpback whale can be assumed to be planar (although in reality this is not the case) and thus conventional techniques such as conformal, affine, or projective transformations can be used. Because the fluke can be rotated in three directions (around the x, y, and z axes) creating perspective distortions, the projective transformation should be the most appropriate method. However, the solution for a projective transformation also requires more control data (at least four well-distributed control points) than the other transformations.

Tests comparing the effectiveness of affine and projective transformations on photographs of a model fluke rotated at a variety of angles generally indicated that an affine transformation accomplished the best overall result because it had more redundant data and the lack of control points below the fluke tips was more likely to produce a weakness in the projective transformation. Further improvements occur when the fluke is divided into four "districts" and separate transformation solutions are used for each of the districts. The fluke is divided into two halves, as often the left and right sides of the fluke define two discrete planes. The central tail stock area also has large amounts of three-dimensional relief (this area appears as a triangular pyramid projecting out from the rest of the fluke); therefore, it is divided into left and right districts to effectively define the other two planar surfaces. As each half of the fluke is transformed separately, only three "useable" control points are available (one fluke tip, the V-notch, and one leading edge base point); essentially eliminating the projective transformation as a possible solution. However, the second leading edge base point is also included as it strengthens the affine solution and provides the fourth control point for tests using projective transformations.

In some images, the two control points at the leading edge of the fluke are covered by water and the operator must estimate the position of these points (and any other control points that cannot be seen). The operator is also required to indicate the strength of each control point (from good to very poor) so that the program is able to modify the search parameters for features measured near less well-defined control points.

First, a primary transformation covering the whole fluke is calculated so that the program can readjust the horizontal position of the two leading edge control points to one-quarter the distance from the center line to the fluke tips to help increase the consistency in the position of these points. The position of a number of minor control points, which are used to define the shape of the fluke, are also displayed on the image. The positions of these points are initially estimated from the transformation parameters and then moved up or down until the edge of the fluke is detected using simple edge-detection techniques (locating the black edge). Often the background ocean contains dark regions, reducing the efficiency of the edge detection technique, in which case the operator can manually reposition the points to accurately define the outline of the fluke (Fig. 2). Most of these points are fixed in position along the horizontal axis (from fluke tip to fluke tip) and the operator only has to move the point up or down to the fluke edge, thus speeding up the process. These points are not used in the transformation process but are used to provide a more accurate shape of the fluke and subdivide the fluke into 18 regions. One additional point is also required (point 20, Fig. 2) to establish the vertical axis down the center of the fluke and the water level on the fluke. A sixth major control point is established by the program that lies halfway between the two base control points. This control point is computed as the intersection between the line joining the two base points and the line from the V-notch to point 20, thus correcting for perspective distortions. This control point, the V-notch point, and one of the leading edge base points are used to compute the affine transformation parameters for each triangular district on the stock area.

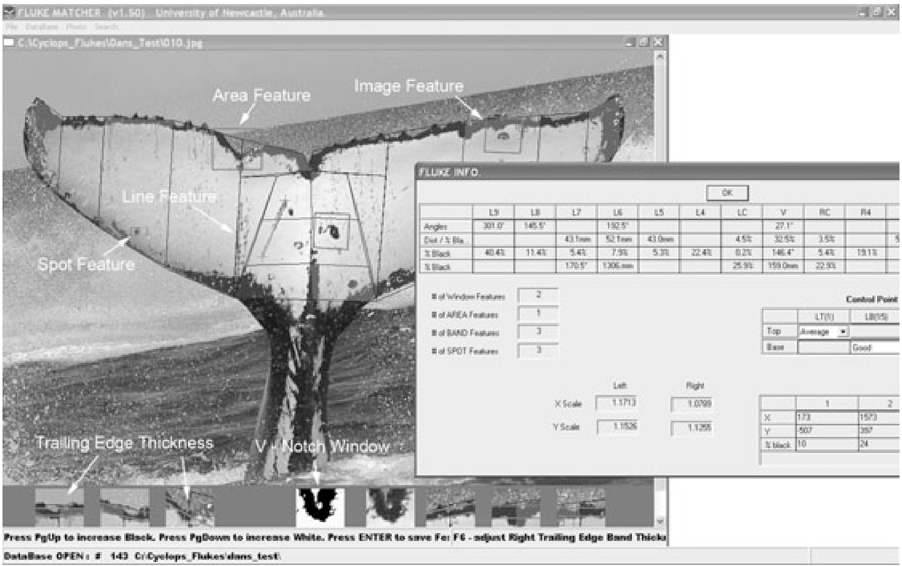

Figure 3. Screen capture of fluke data showing features measured and information box. The "Fluke Info." box shows the values computed for each feature (*e.g.*, "Angles: L9 = 301.0 " is the angle at top of left fluke tip), the quality of the control points, and the number of additional features measured. Trailing edge black bandwidths and shape of pigmentation around the V-notch are shown at the bottom of the diagram. The 18 regions are color shaded wherever black pigmentation is found.

The next stage of the program ("Process Image") re-calculates separate transformations for the four districts but the transformed data are all defined in the one unique reference system. Angles and distances that define the shape of the fluke tips and the center V-notch area are computed, the thickness of the black band along the trailing edge, and the general shape of the fluke. If the photograph has poor contrast, the measured thickness of the black band along the trailing edge may sometimes require correcting by manually adjusting the position of the two lines that identify the top and bottom edges of the trailing edge band. The fluke is divided into 18 regions and the percentage of black pigmentation (black to white ratio or b/w ratio) is calculated for each of these regions, with the areas of black pigmentation detected by the program being shaded on the image (Fig. 3). If required, the operator can manually adjust the b/w ratio in each region by increasing or decreasing threshold levels until the shaded black pigment on the displayed image most closely matches the actual black pigment on the original background photograph (Fig. 3). There are an intensified number of regions around the central area of the fluke, as quite often southern hemisphere humpback whales have distinctive black markings in this area.

Additional features found on the fluke are also measured to provide additional information to help identify the whale. Five types of additional features can be measured (Fig. 4):

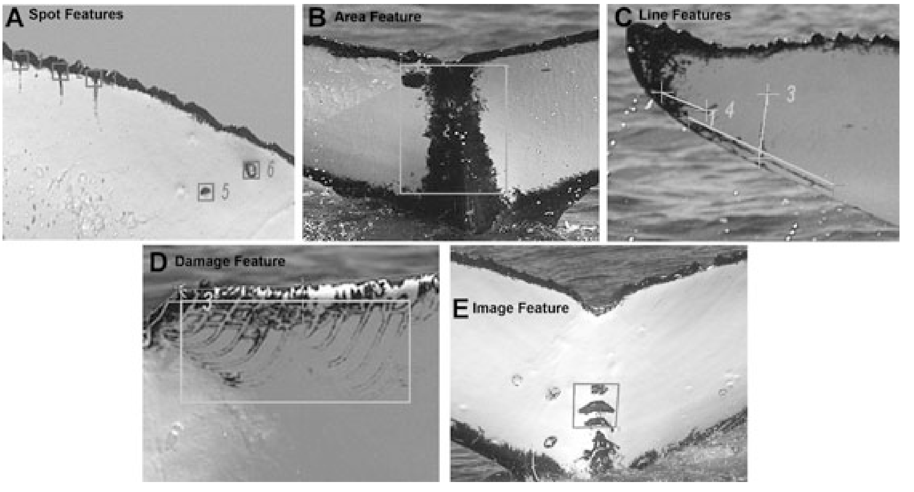

Figure 4. Typical examples of the five types of additional features that can be measured: Spot, Area, Line, Damage, and Image features.

(1) Spot Feature. Small circular features, often ring marks left by barnacles that have dropped off, cookie cutter shark bites, or small pigmentation "bleed" marks near the trailing edge.
(2) Area Feature. Larger area that defines an "area of interest" that may have a number of distinguishing marks.
(3) Line Feature. Band or line that identifies scratch marks *etc.*
(4) Damage Area. Areas of damage, such as missing parts of the fluke or orca rake marks.
(5) Image Feature. A well-defined distinctive mark or pattern that can be recognized by its shape and is suitable for image-matching techniques.

Once the operator is satisfied with the measurement phase, the data can be stored in a database, or a search for the same fluke can be done against images already stored in the database. The major information stored in the database consists of the (x, y) co-ordinates (measured from the base of the V-notch as [0, 0]) and quality of the major and minor control points, the position of any additional features and the value for each, and other general information about the photograph (location, date, photographer, *etc.*).

Because the program uses a wide variety of information about the fluke, this also allows the searching procedure to be flexible. These adaptable matching techniques are automatically set, but in the event that a match is not found toward the top of the rankings the operator can choose to search using different criteria that are potentially more effective at detecting any matches for that particular image by assigning different weightings to the various measured characters for the search (*e.g.,* for flukes that have obvious features or pigmentation patterns). A flow diagram showing the general procedure used in the search module is shown in Figure 5.

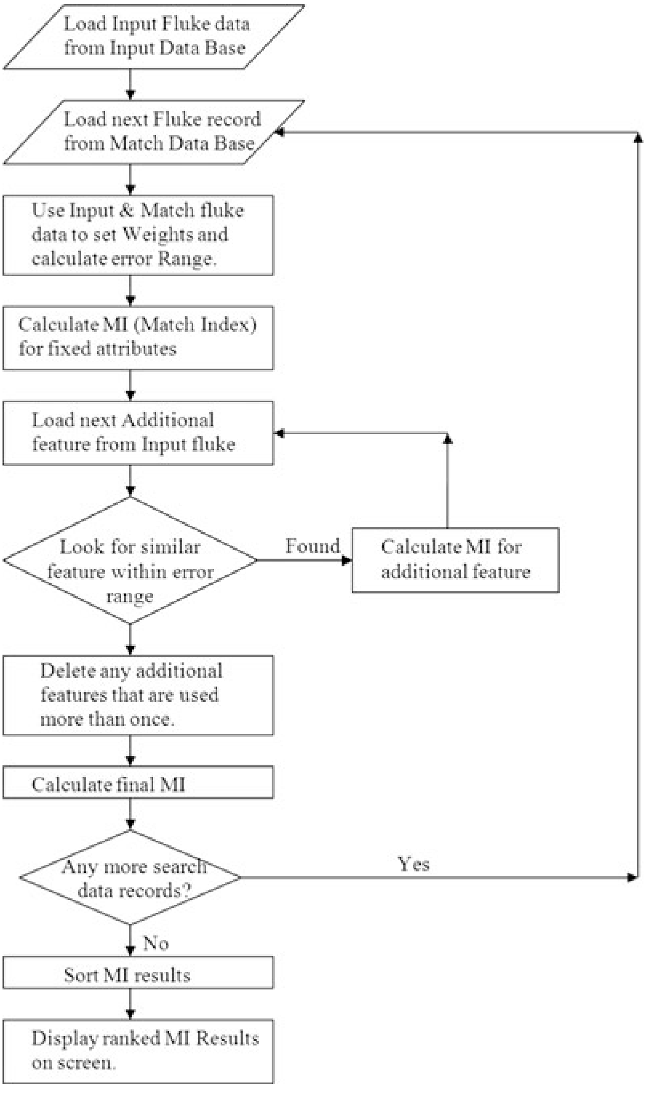

Figure 5. Flow chart for Search phase of Fluke Matcher.

There are a minimum of 32 parameters used in the matching process, plus any additional features measured by the operator, permitting up to a maximum of 61 attributes that can be utilized: (1) 18 regions of b/w ratio, (2) six locations where the thickness of the black band along the trailing edge is measured, (3) three parameters defining the shape and size of the central V-notch, (4) four parameters defining the shape of the fluke tips, (5) five parameters describing the overall shape of the fluke, (6) up to 10 spot features, (7) up to five area and damage features, (8) up to five line features, and (9) up to five image features.

These attributes can be broken into two broad groups: fixed (1–5) and user defined (6–9). Fixed attributes can be directly compared to the same attributes measured on other fluke photographs. For example, the measure of similarity between the thickness of the trailing edge at any of the six defined points in any two fluke images being compared can be computed as:

$$MI_F = TE_1/TE_2 \times 100,$$

where MI_F is the Match Index (0–100) for the feature measurement, TE_1 is the shorter of the two Trailing Edge thicknesses being compared (mm), and TE_2 is the longer of the two Trailing Edge thicknesses being compared (mm).

Other distances, the b/w ratio, and angles can also be compared in a similar way. Some aspects of the user-defined features such as the size, length, or type of feature are also determined in a similar manner.

As the transformed position of each user-defined feature is known, the program only attempts to match features that are found in the same area of the fluke. An indicator of the match strength for each feature is based on how closely the features are located to each other plus the similarity in their measurements (type, size, shape, b/w ratio, *etc.*). The matching process is more flexible for features located near lower quality control points to allow for any weakness in the transformation solution. The distance between the locations of the feature on both fluke images is computed. If this distance is within the allowable error range then Fluke Matcher computes the position match index (MI) for this attribute:

$$MI_P = (R - D)/R \times 100,$$

where MI_P is the MI (0–100) for the feature position, R is the allowable range in the distance measured (mm), and D is the distance between the positions of the feature on the two images (mm).

The value of R was computed empirically and varies from 750 mm for fluke images with little rotation and when the feature is located near good quality control points, to over 1,500 mm for images with large rotations and poor quality control points. The final MI for the user-defined features can then be computed:

$$MI = {}^1/_2 MI_F + {}^1/_2 MI_P.$$

The weight for each feature is initially set to one (1.0) and then adjusted to a value between 0 and 12 based on a number of factors: (1) the quality of the control points near the feature, (2) the amount of rotation in the photograph, (3) the quality of the photograph (contrast, focus *etc.*), (4) the importance of the feature, (5) regions that are hidden by water, and (6) the number of features in each image and the number of features matched.

In addition to this, the operator can utilize a number of different search techniques by selecting the protocol used in the matching process. A number of criteria were tested: (1) standard search, (2) feature search; more weight given to additional features, (3) region search; more weight given to b/w ratio in the 18 regions, (4) trailing edge search; more weight given to trailing edge characteristics, (5) manual search; operator defined weights for different features, (6) bandwidth matching; the use of bandwidth matching techniques when matching the trailing edge, V-notch area, and image features, and (7) image matching; image-matching techniques used to match trailing edge, V-notch area, and image features.

Individual image features, the V-notch area, and the trailing edge can be matched in three different ways. The standard method is to compare the b/w ratio in each of the areas, which is the technique used in the first five (1–5) search methods. Bandwidth matching (6) is a one-dimensional image-matching technique that compares the vertical widths of the black patches in the areas being compared. This technique was especially developed for the program. Image matching (7) uses image-matching techniques to calculate the similarity between image features, shifting the window in two directions until the b/w patterns best match. No rotation or scaling should be required as this is already accomplished in the initial fluke transformation phase.

A data set of 117 humpback whale fluke photographs was used to develop and test the new computer-aided fluke matching system, Fluke Matcher. These photos were collected during vessel surveys off the coast of Ballina and Byron Bay on the east coast of Australia between 2003 and 2006. All photos had previously been manually reconciled to determine any matches. To ensure that poorer quality photographs did not initially influence the test results, all images used in the tests had passed photo quality screening criteria (as used by the SPLASH project and outlined in Calambokidis *et al*. 2008). There were a possible 94 matches within the data set.

During the search process, Fluke Matcher matches each feature measured against each image in the database, giving each a score (0–100) for the probability that the feature properties are identical. In most instances comparing the b/w ratio for image features, the V-notch, and the trailing edge gave the best results (67% of the time) as this method is fairly constant even with poorer quality images; however, sometimes the bandwidth method performed better (33% of the time), especially with images that had little fluke rotation and good contrast. The image-matching technique generally had good results but was rarely significantly better than the other methods and it was much slower to process (40 × slower).

An overall weighted MI(0–100) is then calculated for each image. All images in the database are then ranked in order from the most likely match down to the least likely match and displayed in that order (Fig. 6). The operator can then scan through the list to visually compare images and identify matches. If the operator cannot find a match toward the top of the list, then another matching technique can be used to search for potential matches.

The average time taken to enter each fluke into the database was approximately 4.5 min, but this procedure is only required to be done once for each fluke entered into the database. Initial tests using two different search protocols for a database of 117 flukes are presented in Table 1, showing rankings for the 94 matching flukes.

The standard search resulted in all of the confirmed matches being listed in the top 30 (25%) of the 117 ordered images (Table 1). Eighty-two percent of the matches were ranked first of 117 and the lowest was ranked 26th. In this instance, the operator

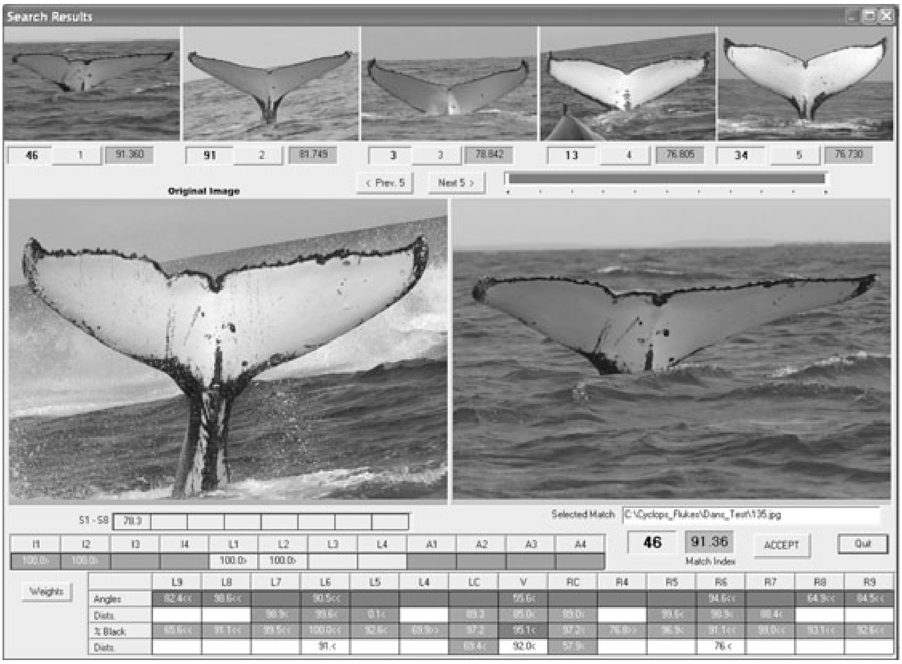

Figure 6. Screen capture of results after a standard search. The top of the diagram lists the match results in order, showing underneath: the photo # (left), the ranking (middle button), and the Match Index (right). The bottom half of the diagram has detailed information on the selected match (#46, rank = 1 in this case). The score for each feature is shown in the grid tables; the symbols <, , >, and are used to indicate the weights applied to each feature. For example, "Angles: L9 = 82.4 " indicates that the score for this feature (angle at top of left fluke tip) of 82.4 has a reduced weight (likely to be caused by the poorer transformation in photo #46).

would only have to scan through the first 22% of the ranked images to find the correct match. Most of the lower ranked matches showed some improvement when other search techniques were used. All matches were ranked in the top 15 (13%) when the best match technique for that particular pair was used. In 89% of the cases, the standard matching technique performed the best (or equal best) and at other times, alternative matching techniques produced marginal improvements.

The matching algorithms used in Fluke Matcher are designed to rearrange the photos so that highly matched flukes occur toward the top of the rankings. Trials showed that using stricter matching protocols would result in more confirmed matches being ranked number one, but also sometimes resulted in correct matches being pushed toward the lower end of the list; generally caused by the results of poorer transformations due to large amounts of rotation in the fluke. Therefore, a more flexible matching protocol provides the best overall results but this may result in larger numbers of similar flukes that are not true matches (false-positives) being ranked toward the top of the potential matches. In the case of a computer-aided matching system such as this, false-positives are of less concern than false-negatives

Table 1. Results of standard and best matching techniques for 117 photographs and 94 possible matches. Lowest ranking result for standard search was 26/117. Lowest result using best search method was 13/117.

Best search types		Standard search results			Best search results		
Search type	Percent of best matches	Rank	# of matches	Percent of matches	Rank	# of matches	Percent of matches
Standard	89.4%	1	77	82%	1	80	85%
Feature	1.1%	1–5	86	91%	1–5	90	96%
Trailing edge	7.4%	1–10	88	94%	1–10	91	97%
Band width	2.1%	1–20	92	98%	1–20	94	100%
Others	0%	1–30	94	100%			

as the operator can simply compare images visually and confirm if each image is a match or not. Increasing the number of additional features could lead to a slight increase in the number of false-positives but more importantly at other times helped avoid false-negatives.

Further testing is now underway with larger data sets. These will be utilized to further refine the system and help design photographic quality standards. In addition, because Fluke Matcher uses a large variety of information, the database record for a particular whale fluke could be compiled from several photographs. For example, if the fluke tips are curled over in one photograph, the four parameters used to define the shape of the fluke tips will be weak and it may be desirable to substitute these values with the equivalent parameters from matched photographs of the same fluke where the tips are not curled. This type of approach can also be used to map changes for some features over time, such as the appearance of new rake marks or other damage to the fluke. This will continuously build more robust database information and therefore improve future search results. The system may also remove some bias that may occur when performing manual matching. For example, a distinctive mark on a fluke may be the primary means of identifying a match when using manual matching techniques but if there are changes to this mark over time then a potential match may be overlooked. This is less likely to occur with the Fluke Matcher system because of the broad range of features and techniques used. Furthermore, Fluke Matcher also removes some of the subjectivity inherent to stratification systems due to the fact that many of the fluke features used by the program are measured automatically rather than entirely relying on human operator decisions. The variety of features used by Fluke Matcher also builds some flexibility into the matching process by allowing for some error in the measurement of any given feature, with the results less likely to be drastically changed compared with simple stratification systems where one error may result in a substantial increase in the probability of missing a match.

The Fluke Matcher system greatly improves efficiency in reconciling humpback whale fluke photo-identification data, thus enabling researchers to use larger data sets in their analyses and increase outputs from photo-identification studies. The use of this computer-based matching system also allows new fluke photographs to be efficiently compared with very large data sets and removes some of the biases involved in comparing data sets based on the expectation of the user as to the likelihood of finding a match.

ACKNOWLEDGMENTS

The authors thank the Australian Government Department of Environment and Heritage for initial funding through the Natural Heritage Trust and the Australian Centre for Applied Marine Mammal Science (ACCAMS) for essential funding support to enable this research to be completed.

LITERATURE CITED

Araabi, B. N., N. Kehtarnavaz, T. Mckinney, G. Hillman and B. Wũsig. 2000. A string matching computer assisted system for dolphin photoidentification. Annals of Biomedical Engineering 28:1269–1279.

Arzoumanian, Z., J. Holmberg and B. Norman. 2005. An astronomical pattern-matching algorithm for computer-aided identification of whale sharks *Rhincodon typus*. Journal of Applied Ecology 42:999–1011.

Beekmans, B. W. P. M., H. Whitehead, R. Huele, L. Steiner and A. G. Steenbeek. 2005. Comparison of two computer-assisted photo-identification methods applied to sperm whales (*Physeter macrocephalus*). Aquatic Mammals 31:243–247.

Caiafa, C. F., A. N. Proto, D. Verganiand and Z. Stanganelli. 2005. Development of individual recognition of female southern elephant seals, *Mirounga leonina*, from Punta Norte Peninsula Vald, applying principal components analysis. Journal of Biogeography 32:1257–1266.

Calambokidis, J., E. A. Falcone, T. J. Quinn, A. M. Burdin, P. J. Clapham, J. K. B. Ford, C. M. Gabriele, R. Leduc, D. Mattila, L. Rojas-Bracho, J. M. Straley, B. M. Taylor, J. R. Urbán, D. Weller, B. H. Witteveen, M. Yamaguchi, A. Bendlin, D. Camacho, K. Flynn, A. Havron, J. Huggins and N. Maloney. 2008. SPLASH: Structure of populations, levels of abundance and status of humpback whales in the North Pacific. Final report for Contract AB133F-03-RP-00078 to U.S. Dept of Commerce. Western Administrative Center, Seattle, WA, U.S.A. 57 pp. Available at http://www.cascadiaresearch.org/SPLASH/SPLASH-contract-Report-May08.pdf.

Chu, K., and S. Nieukirk. 1988. Dorsal fin scars as indicators of age, sex, and social status in humpback whales (*Megaptera novaeangliae*). Canadian Journal of Zoology 66:416–420.

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

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

Katona, S. K., B. Baxter, O. Brazier, S. Kraus, J. Perkins and H. Whitehead. 1979. Identification of humpback whales by fluke photographs. Pages 33–44 *in* H. E. Winn and B. L. Olla, eds. Behavior of marine animals—current perspectives in research. Volume 3. Cetaceans. Plenum Press, New York, NY, U.S.A.

Kehtarnavaz, N., V. Peddigari, C. Chandan, W. Syed, G. Hillman and B. Wursig. 2003. Photo-identification of humpback and gray whales using affine moment invariants. Image Analysis, 13th Scandinavian Conference, SCIA 2003, Halmstad, Sweden, June 29 - July 2, 2003, ProceedingsSeries Lecture Notes in Computer Science, Vol. 2749 Bigun, Josef; Gustavsson, Thomas (Eds.) 2003 Proceedings 2003 2749:109–116.

Lillie, D. G. 1915. Cetacea. British Antarctic ("Terra Nova") expedition, 1910. Natural History Reports of Zoology I:85–124.

Ranguelova, E., M. Huiskes and E. Pauwels. 2004. Towards computer-assisted photo-identification of humpback whales. Proceedings of International Conference on Image Processing 2004. Volume 3:1727–1730.

Schevill, W. E., and R. H. Backus. 1960. Daily patrol of a *Megaptera*. Journal of Mammalogy 41:279–281.

Speed, C. W., M. G. Meekan and C. J. A. Bradshaw. 2007. Spot the match: Wildlife photo-identification using information theory. Frontiers in Zoology 4:1–11.

Tomilin, A. 1957. Cetacea. Mammals of the U.S.S.R. and adjacent countries. Volume 9. English Translation, 1967, Israel Program for Scientific Translations, Jerusalem. 717 pp.

Van Tienhoven, A. M., J. E. Den Hartog, R. A. Reijns and V. M. Peddemors. 2007. A computer-aided program for pattern-matching of natural marks on the spotted ragged-tooth shark *Carcharias taurus*. Journal of Applied Ecology 44:273–280.

Received: 14 April 2009
Accepted: 17 November 2009