Australian Marine Mammal Centre

Final Report

(subclause 9 and Schedule Item 5 of the Funding Agreement)

- **Project No**. – 0809/7
- **Title -** Final development of a new computerised fluke matching system and creation of a fluke database for humpback whales photographed off the east coast of Australia from 1999-2005
- **Chief Investigator/s** – Dr Hendrik (Eric) Kniest, Prof Peter Harrison & Mr Daniel Burns
- **Organisation** – University of Newcastle

Activity Period – 20 April 2009 – 15 May 2010

Table of contents

1. Activity Summary
2. The Outcomes/Objectives
3. Appropriateness
4. Effectiveness
5. Financial Account of the Activity

1. Activity Summary

A clear summary of approximately 500 words outlining the work undertaken and any significant findings (for publication on the Department's web site)
Ongoing development of the computer-aided photo-identification matching system, *Fluke Matcher*, has resulted in a fully functional version of the program that is able to efficiently identify individuals and find resights in photo-identification catalogues of humpback whales. *Fluke Matcher* v3.2.2 is superior to previous versions and is now ready to be distributed and used by researchers.

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, and additional computerised image-matching techniques, to produce an efficient and reliable matching system (Kniest et al., in press). Initial user input is needed to identify control points used for the transformation of the fluke image onto a common reference frame. The system then measures key features of the fluke, including parameters to describe the shape of the fluke, and black and white pigment distribution in different regions of the fluke. The operator then highlights other distinctive features of the fluke that enable identification, such as spots, lines and areas of damage. When searching for a given fluke, the program ranks all of the images in the database from the most likely to the least likely match for that fluke, and the operator then visually checks the ranked images to find the match(es).

A database of 440 flukes was entered into *Fluke Matcher* (FM) and used to develop and refine the program, including improving the data entry phase, identifying which features result in superior and consistent matching results, and identifying the best matching algorithms. Better methods were also introduced to measure the characteristics of the trailing edge and the v-notch area of the fluke, including more information on the shape of these areas.

Tests using a number of different operators (both experienced and inexperienced) were conducted to improve the consistency of the input phase of the program, to refine the data-entry protocols, and to improve search methods. Alternative matching algorithms were also developed to make use of existing information within datasets, so that matches that had already been found in previous searches were utilised to remove duplicates from subsequent searches. A *Fluke Matcher* user manual has also been developed and refined throughout the project.

Several sets of pre-matched fluke images were used to test *Fluke Matcher*. The primary tests were conducted using a set of 860 fluke photographs of 367 whales, with each whale represented by at least two and up to a maximum of five images. An additional 386 images for which there were no matches in the data set were also added to create a very large database of 1247 images (754 whales and 1314 match pairs) with which to test the program.

Results of diagnostic tests showed that by using *Fluke Matcher* (v3.2.2), 48.2% of all matched pairs were found within the top 5 ranked images (i.e. 5/1246 = 0.4%) of the database, and 90.3% of all matched pairs were within the top 120 ranked images (i.e. 120/1246 = 9.6%).

A simulated search utilizing advanced search procedures resulted in 98.5% of all matched pairs being

found in the top 10% of the ranked database, and 99.5% in the top 20%. These results show that the best results for minimal effort are achieved within the top 10% of the ranked search results.

A third data set consisting of 40 images (20 pre-matched pairs) was used to test efficiency when multiple users (two experienced, two neophytes) entered images into *Fluke Matcher*. This test showed that experienced users were able to consistently enter flukes and efficiently find the matching images entered by other experienced users. Variation in consistency when comparing flukes entered by neophytes suggest that some training is necessary for before new users of *Fluke Matcher* are able to achieve optimal results, although one neophyte showed strong agreement with the experienced users almost immediately.

2. The Outcomes/Objectives

The degree to which the Activity has achieved the objectives

Three main objectives were identified to complete the development of *Fluke Matcher*:

1) **Complete the final development of the *Fluke Matcher* computer-based matching system and database to identify and match unique features of humpback whale flukes.**
Like all software, the development of *Fluke Matcher* is an ongoing process, however the current version (3.2.2), is now ready to be distributed and used by researchers. The system is now more user-friendly than previous versions, allowing flukes to be imported and/or deleted from databases, for databases to be stratified by photo quality, and the amount of time needed to input a fluke has been reduced. Previous versions of *Fluke Matcher* only used a 'Standard' search procedure, utilising one of the different types of search methods available and performed the search over the entire database. Existing information, such as already known matches and results from any previous searches were not utilised, but this provided a good method of testing the efficiency of each of the different search methods.

A database of 440 flukes was entered into *Fluke Matcher* and used to accomplish these tasks. Some of the main improvements for version 3.2.2 include:

a) A better method has been adopted to enter detail about the trailing edge and the v-notch area of the fluke, which includes more information about the shape of these areas.
b) Key features have been defined which lead to improved methods for the final search results, and the establishment of better protocols on how to measure these data.

c) Superior matching algorithms have been adopted based on these protocols.

d) The efficiency of the search phase has been improved by using information about known matches. The 'Advanced Search' method improves the efficiency of the search phase by using information about already known matches and the results from any previous search(es). For example a 'standard' search may be carried out and no matches found in, for example, the top 20 ranked images. A second search, possibly using a different search method, may then be conducted, but as the program knows that the 20 images listed at the top of the previous search do not contain a match, then these images are left out of the second search. Both the first and second searches also use information about any known matches, and the program displays only the best-ranked photo from each of these matches.

e) A modified search procedure which puts more emphasis on the trailing edge characteristics and additional features was developed in 2010. This resulted in many more matches being found towards the top of the rankings, although some of the poorer matches did not improve.

f) A Simulated Search protocol for progressively building the match database from an established catalogue has also been developed (see April 2010 in Table 1). The catalogue was re-ordered from best to worst photo quality and a search was then performed for each photo. Only the top 20 ranked images (about 5% of the catalogue of 440 images) were inspected; 598 out of 610 matches were found using this method. The search was then repeated for any un-matched photos, with all remaining matches being found on the second iteration. When the top 40 ranked images were scanned, over 99% of the matches were found in the first iteration. This type of search retains information about previous matches as the simulated search progresses. The search also becomes quicker as the match database is built.

Improvements in the various stages of development for the catalogue of 440 fluke images are shown in Table 1. Photos of varying quality were used in the tests, including images that would not have passed standard photo quality screening criteria (i.e., the SPLASH protocols – see Calambokidis et al. 2001 for further details).

Table 1. Summary of results for tests of match search rankings using a database of 440 images in *Fluke Matcher*.

Search rankings	Nov. 08	Dec. 08	July 09	Feb. 10	April 10*
0 – 5%	457	473	518	544	598
5 – 10%	61	66	52	35	(6)
10 – 15%	23	23	17	12	-
15 – 20%	22	20	5	12	-

20 – 25%	8	6	0	1	-
25 – 30%	3	0	0	0	-
Other	12	0	0	0	-

*Note: Some extra matches were used in 2010 and the April 2010 results used a new matching protocol to build the match database.

2) **Rigorously test the system against a very large fluke catalogue to measure its efficiency and ensure its reliability for use in photo-identification studies.**

Numerous tests of the *Fluke Matcher* system have been conducted and are detailed in Appendix 1. The primary tests used to evaluate the efficiency of *Fluke Matcher* involved a very large database (1247 fluke images of 754 whales) of fluke images entered by operators experienced in the use of the program. A further test involving multiple operators (two experienced, two neophytes) entering flukes into the system was undertaken to test for agreement (or lack thereof) amongst multiple users.

The very large database consisted of 1247 fluke photographs collected during humpback whale surveys at Hervey Bay, Byron Bay and Ballina from 1994 to 2006. Of these, 860 images (representing 367 whales) had at least one and up to four matches in the data set. An additional 387 images for which there was no match in the data set were also included to test how efficiency was affected by the inclusion of unmatched images. This resulted in a total of 1314 pair-wise matches on which to test the efficiency of the program. The photos were of varying quality, from very poor to excellent. A composite, weighted index of the 5 SPLASH scores, the *Fluke Matcher* photo score (FMps), was computed by the program for each image on a scale of 1 to 10 (bigger is better).

The standard search procedure using all 1247 photos resulted in 1194 (91% of the possible 1314) matches listed in the top 10% of the ranked images, and 1263 (96%) listed in the top 20%. 5 matches (0.4%) ranked in the bottom 50% of the database, and were caused primarily by poorer quality photographs, or by the markings on the fluke showing substantial changes over time. Analyses of the effect of photo quality on matching efficiency showed that efficiency increased greatly as FMps increased from 4 to 6, but improvements were relatively small for FMps > 5. The order in which a search was conducted was also found to be important, with rankings being significantly better when using a higher quality image as the 'target' to conduct a search for a 'match' image.

The test was repeated using only the images that had a photo score of 6 or greater, resulting in 94.5% of the total possible matches (reduced to 990) being found in the top 10% of the ranked image list. This small improvement in the results does not justify leaving out a large proportion of the photos, as many

of the poorer quality photos did have very good results; the poorer results tended to occur when the flukes very few distinguishable marks and the detail on the trailing edge was difficult to see.

A Simulated search utilizing the Advanced Search procedure was also carried out. Using all of the 1247 photos in the catalogue, 94.4% of the matched pairs were found when the top 60 (4.8%) of the ranked images were inspected, with 98.5% of the matched pairs found in the top 120 (9.6%), and 99.5% found in the top 240 (20%) (Figure 1). These results indicate that the optimum inspection range is somewhere between the top 5 and 10% of the ranked results, with minimal gain when inspecting beyond this range.

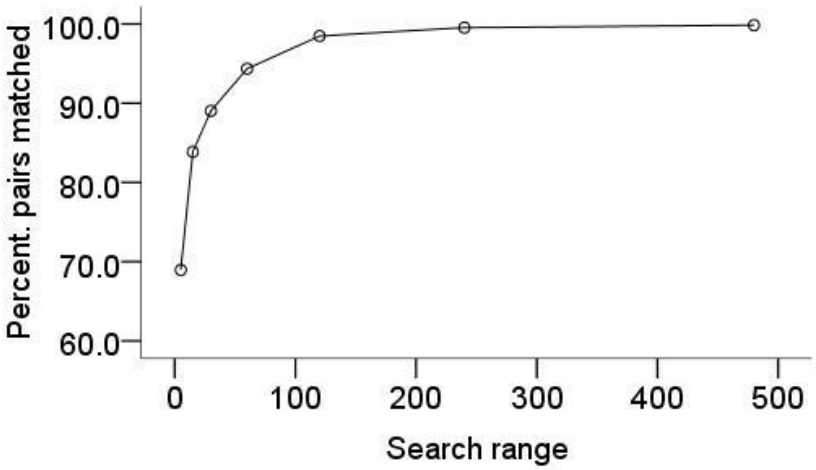

Figure1. Percentage of pairs matched by search range for a data set of 1247 images

It should be noted that the efficiency of the system is likely to increase as the number of repeated sightings of the same individuals increase. This is because the number of implied matches would increase during the search process.

A set of 40 images consisting of pre-matched pairs of 20 individuals was employed for tests of the variation among 4 *Fluke Matcher* users. While all users were experienced in matching humpback

whale fluke images, two were experienced in the operation of the program and two were neophytes. The neophytes were given 1 day's (6 hours) training by the experienced users using a set of images other than those used for the test. Results of the tests were variable. When comparing matching efficiency, the results were significantly better if both users were experienced than if one or both users were neophytes. However, one of the neophytes showed much greater consistency and agreement with the experienced users than the other neophyte (see Appendix 1 for more detail). These tests suggested that it is very important for users to have some experience with the system before consistent results may be achieved, and therefore training workshops and/or the development of online tutorials will be investigated.

3) **Develop a User Manual for the *Fluke Matcher* system, and develop and test a new set of photo quality screening criteria appropriate for the system.**
A User manual has been developed for *Fluke Matcher*, outlining the protocols for creating a fluke catalogue database, entering images into the database, and conducting searches within the system (Appendix 2).

The Cascadia Research Collective Fluke Screening criteria, also known as the 'SPLASH protocols' (outlined in Calambokidis et al. 2001) have been adapted for use in *Fluke Matcher*. When saving a fluke into the database, the program estimates the SPLASH scores for the image (including a score for each of the five quality criteria – fluke angle, lateral angle, proportion visible, exposure / contrast, and focus), based on the characteristics of the entered data. The operator can also change the SPLASH scores manually if required. *Fluke Matcher* then uses the SPLASH scores to calculate an overall photo quality score (ranging from 0 to 10). The photo score is weighted higher on the fluke angle and rotation angle as these play an important part in the transformation process. This overall *Fluke Matcher* photo score provides a good indicator on the likely success of finding a match in the top rankings. Matching efficiency increases greatly as FMps increases from 4 to 6 but improvements are relatively small for FMps > 5.

3. Appropriateness

The appropriateness of the approaches used in the development and implementation of the Activity
The approach used to develop and test *Fluke Matcher* has proven to be highly appropriate, resulting in a substantial increase in the capability and efficiency of researchers to analyse humpback whale photo-identification data. The program measures 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 individual identification. Additional computerised image-matching techniques are also used to help produce a reliable matching system.

Fluke Matcher does not rely on pure image matching techniques, but instead aims to substantially improve efficiency in manual matching methods. The program greatly reduces the time needed to identify individuals and find resights in photo-identification catalogues of humpback whales. Extensive diagnostic tests of the system have been used to develop and improve the program, and the most recent tests (outlined in detail in Appendix 1) show that *Fluke Matcher* is now ready for distribution and use by other researchers.

As *Fluke Matcher* uses a wide range of criteria, based on multiple key features of humpback whale flukes that are normally utilised by manual matching methods, researchers can quickly familiarise themselves with the program.

The program is based on the unique characteristics of southern hemisphere humpback whale flukes, but could be readily adapted to suit northern hemisphere whales. The program also has the potential to be adapted for other marine mammal species.

4. Effectiveness

The degree to which the Activity has effectively met its stated objectives

The major objectives of this project have been met effectively. Fluke Matcher has been developed to a point where the current version of the program (v3.2.2) is now ready for distribution. Compared with previous versions of the program, v3.2.2 is more user-friendly and procedures within the program are more efficient. A new method of inputting the trailing edge and v-notch information is now quicker, easier and more consistent. Analyses of tests have led to current best protocols being adopted with respect to data entry and search methods in the program. The time taken to enter a new fluke into the database is around 4-6 minutes (depending on the quality of the photograph and the number of features that must be entered).

Rigorous tests of *Fluke Matcher* using a very large database show that it greatly increases matching efficiency for analyses of humpback whale photo-identification data (outlined in detail in Appendix 1). For tests using optimal search techniques, 98.5% of all matches were found in the top 10% of the catalogue in *Fluke Matcher*. This demonstrates that the program represents a reliable and efficient

system for photo-identification researchers.

A user manual has also been developed for *Fluke Matcher* (Appendix 2), and new photo quality screening criteria have been developed and tested.

The system improves on visual manual matching in a number of ways:
1) Each fluke needs only to be entered into the system once (the image does not have to be rescanned, categorised or stratified each time a match is undertaken).
2) The search process significantly increases matching efficiency by reducing the total amount of images that needed to be compared and reduces some of the bias of manual matching.
3) All information (including fluke data, photo data and match data) is stored in a database structure that can be easily accessed and cross-referenced by researchers. The match data output file is also easily adapted to create a capture history file of all whales in the database that can then be used for capture-recapture analyses.
4) Any new image entered into the system can be quickly matched against any of the existing databases in the system. The time taken to search for a match in a database of about 1000 images is a matter of seconds. It may then take the operator a few minutes to scan through the ordered list to find any potential matches. This represents a highly significant increase in the efficiency of matching humpback whale flukes, especially in very large catalogues.
5) The new method for building the match database further improves the efficiency of finding potential matches and allows for a means of quickly finding a match with a high probability outcome.

A published paper outlining the development and features of the Fluke Matcher system is currently in press and is included in Appendix 3:

Kniest, E., Burns, D. and Harrison, P. (in press) 'Fluke Matcher': A computer-aided matching system for humpback whale (*Megaptera novaeangliae*) flukes. *Marine Mammal Science* 26(3).

A second paper describing tests of the efficiency and reliability of *Fluke Matcher* is in preparation.