

A note on an automated system for matching the callosity patterns on aerial photographs of southern right whales¹

LEX HIBY* AND PHIL LOVELL†

Contact e-mail: lex@conres.demon.co.uk

ABSTRACT

Photographs showing the callosity patterns of southern right whales (*Eubalaena australis*) are currently compared by eye to identify individuals and monitor their occurrence within certain areas. This paper describes software designed to reduce the number of by eye comparisons required to maintain each of the existing local photo-identification catalogues. The software is used to extract, from each photograph, a viewpoint-independent description of the shape and location of each callosity which generates a parallel catalogue of extracts. This is then compared with the description extracted from each new photograph to generate a list of similarity scores and thus highlight likely matches. The software can also be used to compare the different catalogues of extracts with each other. Using a test set of 67 photographs of 23 whales taken from 1974 to 1986, the software reduced the number of by eye comparisons required to identify all individuals by 93% when compared with a purely random search.

KEYWORDS: PHOTO-ID; SOUTHERN HEMISPHERE; RIGHT WHALE

INTRODUCTION

A study of the feasibility of automated matching for right whale photographs, commissioned by the International Whaling Commission, was completed in 1990 (Lovell and Hiby, 1989). In 1997, the International Fund for Animal Welfare provided funds to update and install such a system for routine use in management of a photo-identification catalogue. A catalogue of aerial photographs of southern right whales near Península Valdés, Argentina, has been maintained at the Whale Conservation Institute (WCI) in Salt Lake City, Utah, since 1971. It holds photographs taken from 1970 to 1990, with those taken since 1991 awaiting entry. The automated system was installed at the WCI in October 1997 and is being used to bring the catalogue up to date. Inter-year similarity scores for photographs of the 1,077 individuals already identified will, in time, provide an extensive performance test for the new system. Currently, the only available test data consists of a sample of photographs of 23 whales taken over the period 1974 to 1986. Brief results are presented below, along with an outline of the system as installed at the WCI.

The most important features permitting individual identification of southern right whales in the WCI and other catalogues (at the University of Pretoria, for example) are the shape and locations of 'callosities', which are patches of thickened cornified epidermis on the top and side of the head. The callosities are grey but are usually covered by dense populations of cyamids that make them look white in contrast to the black skin. Fig. 1 (reproduced from Payne *et al.*, 1983) shows a typical distribution of callosities. The 'bonnet', 'coaming' and 'eyebrows' are always present but the shape of the bonnet and the number and location of the 'rostral islands' and 'lip patches' vary; this allows most individuals to be distinguished and given individual identification numbers.

The automated system makes use of the variation in shape, number and location of callosities on the rostrum and lips. The objective is not to replace visual comparison of the

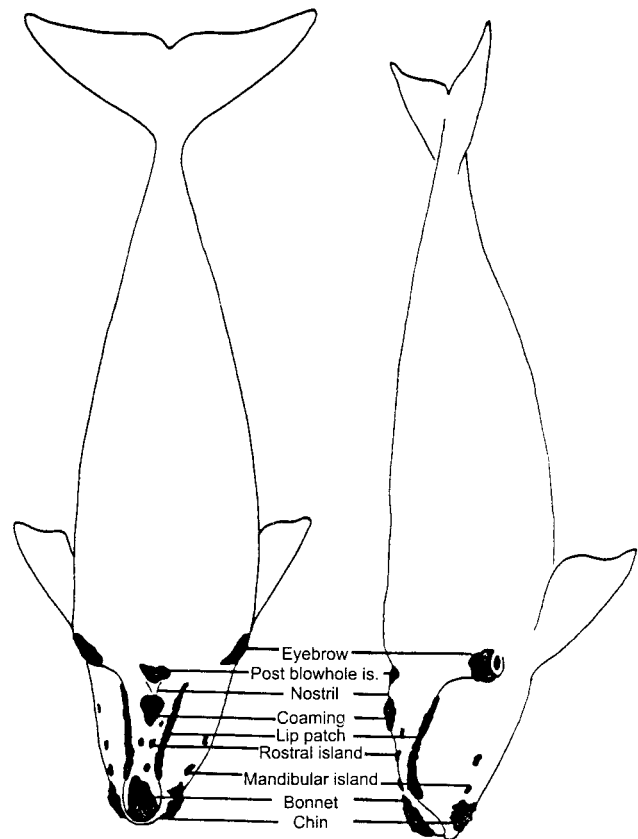


Fig. 1. Typical distribution of right whale callosities (reproduced from Payne *et al.*, 1983)

callosity patterns but to suggest the order in which the existing library be searched for matches to a new photograph, or to eliminate from the search those photographs in the library that cannot possibly be potential matches of the new one. The method used is to extract from each new photograph a simplified map of the callosity pattern; this is used to generate similarity scores with extracts from photographs already in the catalogue. The list

¹ Presented to the IWC Scientific Committee as SC/M98/RW38.

* Conservation Research Ltd., 110 Hinton Way, Great Shelford, Cambridge CB2 5AL, UK.

† Sea Mammal Research Unit, Gatty Marine Laboratory, University of St. Andrews, St. Andrews, Fife KY16 8LB, UK.

of scores is then used to order the search and/or eliminate the least likely photographs. Thus, two main programs are required, one to extract the map in such a way that the result does not depend on the viewpoint of the camera, the other to calculate the similarity scores.

DESCRIPTION OF THE SYSTEM

The first program was adapted from that developed for grey seal pelage patterns (Hiby and Lovell, 1990), which has been used since 1991 to monitor local grey seal populations in the North, Baltic and Irish Seas. A three-dimensional (3D) surface model of the head is used which allows for differences in viewpoint. Each new photograph is scanned into a computer and the digital image displayed on a monitor. The 3D surface model is superimposed on the image, the orientation and scaling of the model being matched to that of the whale in the photograph by identifying four points on the image with the mouse cursor. The points used are at the front margin of the bonnet, the rear margin of the coaming and the upper margin of the eyebrow callosities. Because those points have specific 3D locations on the model, their screen locations should define the required rotations and scaling. The distribution of callosities over the surface of the rostrum and lips is then extracted from the region of the image underlying the appropriate section of the 3D model. Fig. 2 shows the bonnet, coaming and eyebrow points marked on the image, the 'best-fit' projections of the corresponding 3D model points and the resulting orientation of the surface model section.

The distribution of callosities is sampled using a raster scan over the model section, so the spacing of the corresponding sample points on the screen varies with the shape of the model surface, being more closely spaced in areas where the surface meets the camera axis at an oblique angle. The resulting extract is a matrix of sample points with each point classified as 'callosity' or 'skin' depending on the value of the image at the location corresponding to that point in the raster scan.

In the grey seal system there is no pre-processing of the digital image before the pelage pattern is sampled; the right whale system, however, requires an operator to interpret the callosity patterns before sampling. This is because areas of pale skin, water splashes and highlights can be easily confused with the callosity patterns, which are themselves variable in appearance due to the growth of cyamids. It would not be sufficiently reliable to have the computer classify pixels as 'callosity' or 'skin' purely on the basis of, say, brightness as compared with a base value. To overcome this problem, the operator is asked to identify samples of callosity and skin from the digital image, which the program then uses to generate a discriminant function based on the colour, saturation and intensity of each pixel in the samples. That function is then used to classify all the remaining pixels in the image, giving an orange shade to those classified as 'callosity' to allow the operator to see the extent of the callosity regions identified. If necessary, the operator can then use the mouse cursor to extend or reduce those regions. Finally, the operator identifies which of the callosities lie along the lip margins so that the entire image is classified

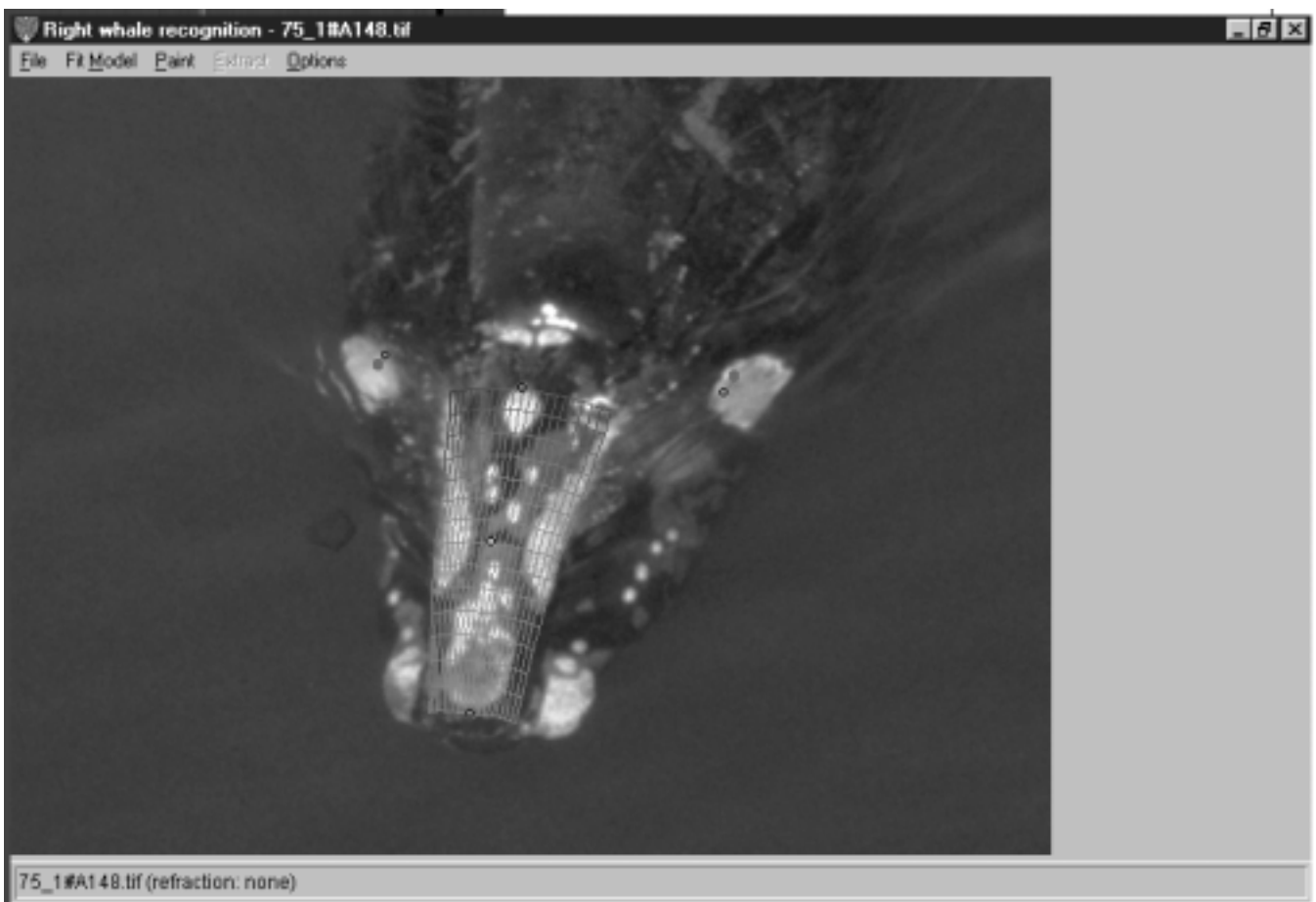


Fig. 2. An aerial photograph of a right whale, taken off Peninsula Valdes in Argentina, displayed on a monitor screen with a section of the 3D surface model superimposed over the rostrum and lip margins. The four dots on the bonnet, coaming and eyebrow callosities are the points marked on the image to define the rotation and scaling of the model required to match its orientation and size to that of the whale in the photograph.

into areas of rostral callosity, lip callosity and skin. Areas of uncertainty can also be highlighted using the mouse cursor. Finally, as the raster scan is performed, any areas obscured by the height of the callosities and the angle of the rostrum to the camera axis are automatically identified as additional areas of uncertainty. Thus, three further types of area are added to the description of the callosity pattern, i.e. possible rostral callosity, possible lip callosity and possible callosity of either type.

This process can take several minutes to complete for each new photograph entered into the catalogue. However, it is only performed once for each photograph. A manual search of a catalogue would require each pattern to be re-interpreted when comparing every pair of photographs, this effort increasing as the square of the number of entries. The key to this digital system is the capture of the operator's best efforts at interpretation; the potential performance of the system is therefore dependant on the skill and experience of the operator.

The second program is designed to calculate similarity scores for any pair of pattern extracts. The 'similarity score' is a measure of total distance from each sample point on the margin of each contiguous region of definite callosity on one extract, to the nearest point of definite or possible callosity of the same type (rostral or lip) on the other extract. The sum distance over all sample points is minimised with respect to the length of the second extract and its rotation about an 'origin' corresponding to the location of the coaming. These distortions can bring regions of definite or possible callosity on the second extract closer to points of definite callosity on the first one. The most similar extracts are thus those with the smallest sum distance score; the minimisation with respect to length and rotation counteracts any increase in the sum distance resulting from a poor fit of the 3D model. The coaming was chosen as the origin because its location on the photograph is the least affected by camera axis orientation and refraction by waves. Because some callosity patterns are sparser than others, all scores are standardised with respect to the distribution of scores across all extracts in the library. The worse (i.e. larger) of the two resulting standardised distance scores is chosen because, if two extracts are from the same whale, the distance score between them will be small when compared with the distance score each would achieve with extracts in the catalogue from other whales.

'Areas of uncertainty' highlighted by the operator or identified automatically by the program are considered as regions of 'possible callosity' during the comparison and therefore tend to reduce the size of the distance score. The measure used does not match pairs of photographs where one member has no callosities in the region of the rostrum or lip margins where the other member has one or more definite callosities. Even a small callosity may give rise to quite a large distance score if the other extract has no definite or possible callosities in or near that location. The measure is weighted in this way because although the exact size and shape of each callosity might be difficult to determine, callosities should not be completely invisible on an acceptable photograph.

The system uses filenames to record decisions about which photographs are of the same whales and stores the summary statistics used to derive standardised scores within the extract files themselves. This allows it to be used with catalogues that have been established over a number of years, where the results of visual comparisons and associated data already form an extensive database. The filename system can exist in parallel to such a database and hence avoids the need to reorganise the original catalogue. A set of

Microsoft Excel modules has been written to facilitate the selection and editing of filenames, and to display the images of the pairings from which the high-scoring extracts were obtained.

The extracts are initially stored as files in a 'pending' directory and then automatically compared, either singly or in batches, to each extract file in a 'library' directory. The standardised distance scores are stored, in ascending order, in result files, one for each selected pending file. Following automated comparison, the selected pending files are moved to the library. The digital images from which the extracts with the best scores were obtained can then be displayed next to the image corresponding to the pending file, to confirm or reject a match. Extract files are assigned unique names (based on whatever system was used to identify the photograph) but the filename can also be extended using an individual identification or 'whale' number. The name of each extract thus identifies which photograph it came from and which whale was in that photograph. Initially, the whale number may be a temporary number used to identify photographs of the same whale in the pending batch, but when a match is confirmed the whale number part of the pending filename is edited to that of the library one. Furthermore, if two files in the library with different whale numbers are found to represent the same whale (because both match to the same pending file) the whale number part of the later filename is edited to that of the earlier one. Thus, the capture history of a given whale is represented in the sequence of filenames.

The whale number part of the filename is also used to collate the similarity scores. In order to minimise the risk of missing a match, different images of the same whale are allowed to build up in the library. When a new file in the pending directory is compared with the library, only the maximum similarity score for each group of library files having the same whale number is chosen for storage in the results file. This avoids the operator having to make visual comparisons with each photograph of a set already known to be of the same whale.

TRAINING AND TESTING

In regular use, the files selected from the pending directory will be for unidentified whales. However, where a backlog of photographs of identified whales is to be entered, as in the case of the WCI catalogue, the filenames for the files in the pending directory will already have established whale number extensions. In that case the similarity scores for any files from the same whale already transferred to the library will be highlighted in red in the results. The performance of the system can thus be monitored during entry of the backlog and any problems, e.g. in interpretation of the patterns before sampling, can be identified.

The system also allows extracts to be displayed for visual comparison (Fig. 3). This allows problems with low-scoring pairs from the same whale to be identified, for example, one or more small callosities might have been missed. The code used to display the extracts also allows two or more extracts to be combined into a new representation. This option is provided for instances where a sequence of photographs may reveal the callosity pattern over the entire rostrum and lip margins, but where on each individual photograph part of the area is obscured. A mosaic of the partial patterns can then be constructed and may be more informative than the collection of partial samples would be. For example, a sample from a different whale might score quite well against each of the partial samples individually but score badly against the

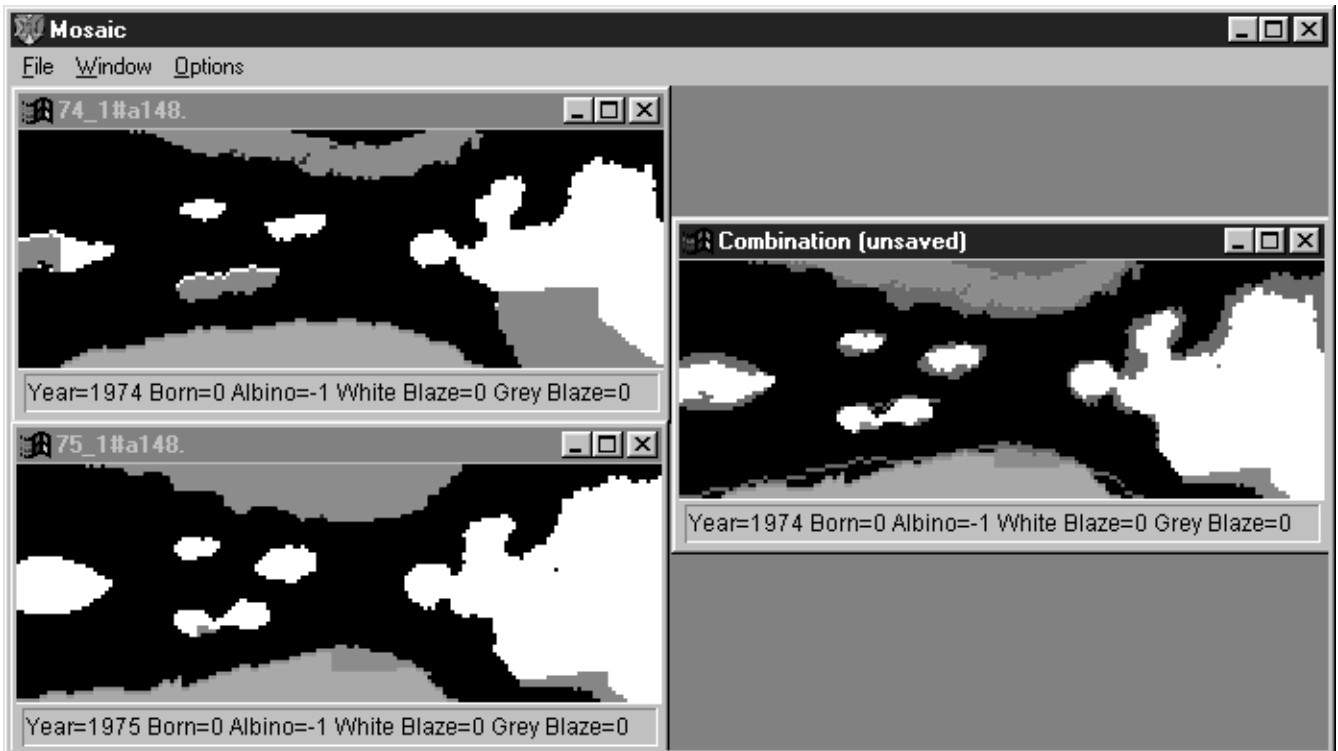


Fig. 3. Extracts from photographs of whale number 148 taken in 1974 and 1975 are displayed for visual comparison, to the left of the screen. The coaming and bonnet callosities are to the left and right of the extracts. The 1975 extract shows three definite rostral callosities whereas in the 1974 extract two of the three have been marked as 'definite' and one as 'possible'. This may be because the operator was uncertain whether a pale area was a callosity or another feature such as a scar or highlight due to the oblique angle of the photograph. The program automatically marks as 'possible' those regions where the extent of the callosity 'footprint' on the rostral surface is uncertain because it is hidden by the height of the callosity. Regions of the bonnet, coaming and right lip callosities have been marked in the same way. The extract to the right of the figure shows that different extracts can be combined if the resulting extract provides a better representation of the callosity distribution.

mosaic because of the relative positions of the callosities over the whole area. To facilitate correct orientation for the components of the mosaic the operator can identify any distinctive feature visible on all photographs in the sequence. When the 3D model has been orientated to the first photograph in the sequence the screen location of that feature is marked. The 3D location for that feature is then added to the locations of the bonnet, coaming and eyebrow callosities and used to aid orientation of the 3D model to the subsequent photographs of the sequence.

The similarity scores generated by the system can be used in different ways. Where most new photographs are of whales that will probably not be represented in the catalogue, it may be best to inspect only those library photographs which achieve a similarity score above a set threshold level. The distribution of scores between matching photographs in the backlog can be used to assess the risk that none of the extracts from matching library photographs will exceed the threshold score. Similarly, the distribution of scores between photographs of different whales will show what proportion fall below the threshold and will therefore not require comparison by eye. Alternatively, if the proportion of 'new' whales among each batch of new photographs is generally small, as is the case in the WCI catalogue, it may be better to search the entire library visually, in the order of descending similarity scores, instead of imposing a threshold. The likely saving offered by the system in this instance is then given by the mean ranking of similarity scores for pattern pairs from the same whale. The results for the small sample of test photographs mentioned in the introduction are presented below in this format.

RESULTS

A sample of 67 colour transparencies was selected by the WCI showing 23 whales over the period 1974-86. Each whale was present in two or three of the years. The sample was selected to represent the range of problems encountered in interpreting the callosity patterns so the average photographic quality over the sample was less good than over the catalogue as a whole. A pattern sample was selected from each photograph, the resulting files copied to the pending and library directories and all the pending files compared, in a single batch, with the library. With the 67 library files collated into the 23 animal groups and the pattern pairs inspected in descending order of similarity score, a total of 54 potential matches had to be eliminated by eye before all the matching whales were located. Had the files been inspected in random order the expected number eliminated by eye would have totalled $67 \times 23/2$, i.e. 771. Thus, about 93% of the matching effort was eliminated by using the automated system to order the potential matches, as compared with a purely random search. The calculation of random search effort is intended only to provide a convenient baseline - actual catalogue searches are never performed in a totally random order because some classification by type is used when patterns are stored in the catalogue.

The histogram (Fig. 4) shows that when the library files were not collated, so that similarity scores were returned for each sample pair, a number of same-whale pairs achieved poor similarity scores. The results indicate that for whales represented by a single pattern sample in the library, the

reduction in search effort provided by the system would be only around 75%. Visual comparison of the low-scoring sample pairs showed that, in every case, the problem was misinterpretation of the callosity pattern, resulting in one or more callosities being missed for one member of the pair. The number of errors would clearly have been reduced if the patterns had been interpreted by someone with more experience of looking at right whale callosities but it is not

possible to say by how much. It should be possible to provide a better estimate once the backlog of WCI photographs has been entered into the library and each pairing subjected to comparison.

ACKNOWLEDGEMENTS

We would like to thank Vicky Rowntree and Roger Payne of the Whale Conservation Institute for the opportunity to develop and test the automated system on the Argentinean right whale catalogue and for permission to use photographs and figures in this paper. We would also like to thank the International Whaling Commission and the International Fund for Animal Welfare for funding the development of the automated system.

REFERENCES

Hiby, A.R. and Lovell, P. 1990. Computer aided matching of natural markings: a prototype system for grey seals. *Rep. int. Whal. Commn* (special issue) 12:57-61.
 Lovell, P. and Hiby, A.R. 1989. Automated photo-identification of right whales, progress report. Paper SC/41/PS16 presented to the IWC Scientific Committee, May 1989 (unpublished). 5pp. [Available from the Office of this Journal].
 Payne, R., Brazier, O., Dorsey, E.M., Perkins, J.S., Rowntree, V.J. and Titus, A. 1983. External features in southern right whales (*Eubalaena australis*) and their use in identifying individuals. pp. 371-445. In: R. Payne (ed.) *Communication and Behavior of Whales*. AAAS Selected Symposia Series 76. Westview Press, Colorado. xii+643pp.

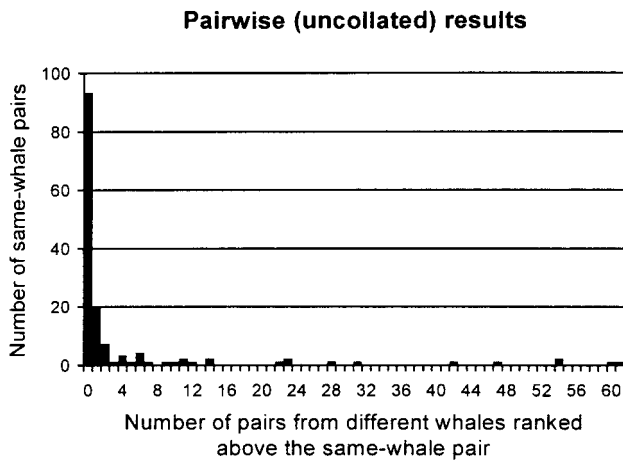


Fig. 4. The frequency distribution for the number of non-matching extract pairs ranked above each matching pair by the software designed to generate pair-wise similarity scores.