ARCTIC
VOL. 64, NO. 3 (SEPTEMBER 2011) P. 342–352

Computer-Assisted Photo-Identification of Narwhals

MARIE AUGER-MÉTHÉ,[1,2] MARIANNE MARCOUX[3] and HAL WHITEHEAD[1]

(*Received 7 September 2010; accepted in revised form 3 March 2011*)

ABSTRACT. Although the narwhal (*Monodon monoceros*) is economically and culturally important to northern residents, sound management of this species is impaired by large gaps in knowledge. Research on this species has been limited partly by the cost of the methods used, and partly because some of these methods are invasive and therefore condemned by Inuit communities. Photo-identification, a non-invasive, inexpensive, and easy-to-use method recently developed for narwhals, uses photographs of natural marks to identify individuals. Its main drawback is the extended time required to process photographs. We developed a computer program to accelerate the identification process and thus mitigate the main drawback of photo-identification. This program uses the locations of notches on the dorsal ridge to compare a new image to each individual in a catalogue and lists those individuals in decreasing order of similarity. We tested consistency in user assignment of dorsal ridge features and the accuracy of the program by comparing sets of known individuals. While assignment errors were common, the program ranked the true match within the first 10% of the catalogue 78% of the time. The program accelerates the matching process by 1.2 to 4.1 times for catalogues ranging in size from 40 to 500 individuals, and the degree of acceleration increases with the size of the catalogue. This program could also be applied to the beluga whale (*Delphinapterus leucas*), another important northern species.

Key words: Arctic, individual identification, *Monodon monoceros*, narwhal, non-invasive methods, photo-identification, software

RÉSUMÉ. Bien que le narval (*Monodon monoceros*) soit une espèce exploitée d'importance économique et culturelle pour les résidents du Nord, la gestion efficace de cette espèce est affaiblie par des lacunes importantes en matière de connaissance de l'espèce. La quantité de recherche sur le narval est limitée par le coût des méthodes utilisées et par le fait que certaines de ces méthodes sont invasives, ce qui est désapprouvé par les communautés inuites. La photo-identification, soit une méthode non-invasive, peu coûteuse et facile d'utilisation, a été récemment mise au point pour le narval. Cette méthode utilise des photographies de marques naturelles pour identifier les individus. Toutefois, le plus grand défaut de cette méthode est qu'elle requiert beaucoup de temps pour comparer les photographies. Nous avons mis au point un programme informatique dans le but d'accélérer le processus d'identification et donc remédier au principal inconvénient de la photo-identification. Ce programme utilise l'emplacement des entailles dans la crête dorsale des narvals pour comparer une nouvelle image à celles d'un catalogue et les ordonne en ordre décroissant selon leur similarité. Nous avons testé la constance de l'utilisateur lorsqu'il attribue les caractéristiques de la crête dorsale et l'exactitude du programme en comparant des photographies d'individus précédemment identifiés. Bien que les erreurs de l'utilisateur soient fréquentes, le programme classe le bon individu parmi le premier 10 % des individus du catalogue, et ce 78 % du temps. Ce nouveau programme permet alors un meilleur rendement du processus d'identification de 1,2 à 4,1 fois plus rapide que sans l'assistance d'un programme pour un catalogue comprenant de 40 à 500 individus. Par ailleurs, plus le catalogue est grand, plus le degré d'accélération augmente. Ce programme informatique pourrait aussi être appliqué au béluga (*Delphinapterus leucas*), une autre espèce d'importance pour les résidents du Nord.

Mots clés : Arctique, identification des individus, *Monodon monoceros*, narval, méthode non-invasive, photo-identification, programme informatique

Révisé pour la revue *Arctic* par Nicole Giguère.

INTRODUCTION

The narwhal (*Monodon monoceros*), a species of economic and cultural importance to Inuit communities, was assessed as "of special concern" and "near-threatened" by national and international conservation agencies largely because of data deficiencies and uncertainties in abundance estimates and trends (COSEWIC, 2004; Jefferson et al., 2009). Although current monitoring methods have made significant contributions to narwhal research, most (e.g., aerial surveys) are expensive, and some (e.g., satellite telemetry) are invasive and therefore condemned by some

[1] Department of Biology, Life Science Centre, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia B3H 4J1, Canada
[2] Current address: Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada; corresponding author: marie.auger-methe@ualberta.ca
[3] Natural Resource Sciences, McGill University, 21 111 Lakeshore, Ste-Anne-de-Bellevue, Québec H9X 3V9, Canada

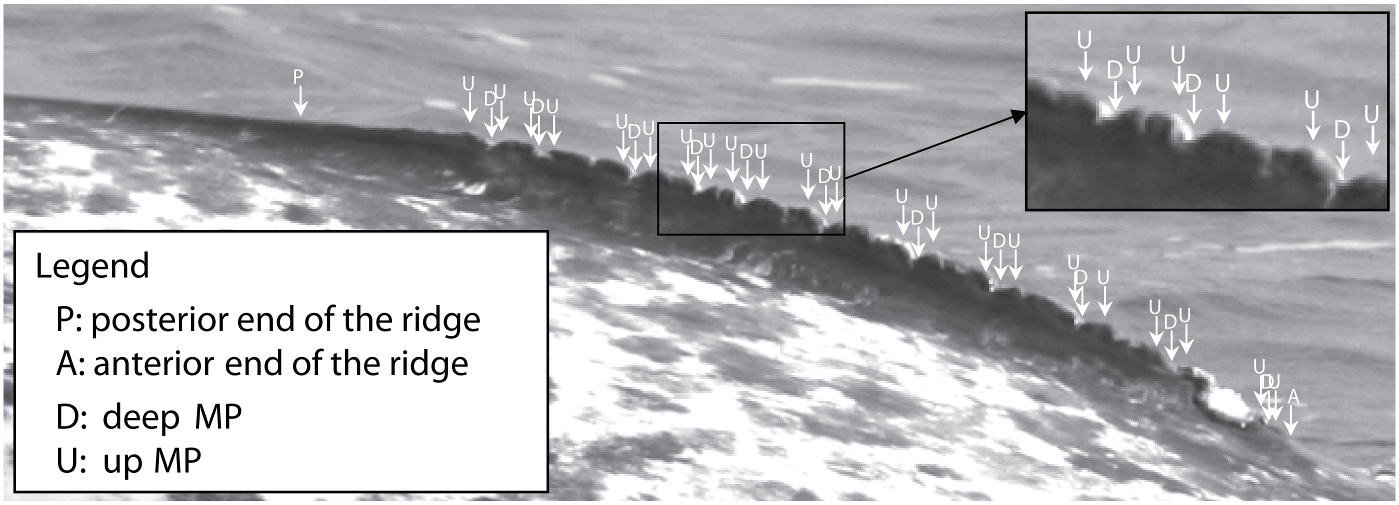

FIG. 1. Example of the mark points (MPs) located on an individual narwhal. The deep and up MPs represent features of the notches, which are indentations that cut the ridge through its entire depth (Auger-Méthé et al., 2010). The anterior and posterior ends of the ridge are also represented.

Inuit communities (George, 2006; Siku News, 2007). These limitations restrict the research conducted on narwhals. As part of an effort to develop non-invasive, inexpensive, and easy-to-use methods that would allow increased monitoring of this species (see also Marcoux et al., 2009), we developed a photo-identification method for narwhals (Auger-Méthé et al., 2010). Photo-identification uses photographs of scars or pigmentation patterns to identify individuals. This technique could be applied to investigate many aspects of narwhal ecology, including abundance estimates and trends, survival rate, and social structure.

Although photographing narwhals is non-invasive and inexpensive, processing photographs is a time-consuming task. For example, comparing a single photograph to a catalogue of a few hundred individuals can easily require an hour of effort. A number of computer programs have been developed to accelerate the identification process for other species (e.g., Hiby and Lovell, 1990; Hillman et al., 2003; Arzoumanian et al., 2005; Gamble et al., 2008). These programs extract information from a photograph entered and estimate how similar this information is to that of each individual previously entered into the catalogue. These programs differ most significantly in the type of algorithm used to calculate similarity coefficients. The choice of the algorithm is highly dependent on the type of marks used for the photo-identification of the species of interest.

Notches of the dorsal ridge appear to be the most suitable marks for identifying individual narwhals (Auger-Méthé et al., 2010). Notches are indentations that cut the dorsal ridge through its entire depth. They are found on 91% of individuals, appear to be relatively stable over time, and are sufficiently variable in shape, size, and numbers to differentiate between individuals (Auger-Méthé et al., 2010). Many programs (e.g., Whitehead, 1990; Hillman et al., 2003) identify individuals by the presence or shape of similar marks, which in other species are found on the dorsal fin or on the fluke. However, none of these programs are directly applicable to narwhals, in part because many rely on landmark features (e.g., the tip of the dorsal fin) to scale the notches' characteristics. The main difficulty with matching narwhals is that the landmarks that can be used to scale the location of the notches, the ends of the dorsal ridge, are particularly hard to locate.

We have developed a matching program for narwhals that is designed to cope with the hard-to-locate landmarks of their dorsal ridge. Here we describe how the program works, emphasizing the extraction of information from the photograph and the algorithm used to rank possible matches. We also evaluate user consistency in locating the dorsal ridge features and discuss the speed, accuracy, and optimization of the program.

MATERIALS AND METHODS

Matching Routine

We chose to base our program on the principles of the matching routine of Whitehead (1990). We wrote our new routine using MATLAB 6.5 (MathWorks, 2007) and its imaging and database toolboxes. The matching routine interacts with Microsoft Access databases, one of which, called the catalogue, contains information on previously matched individuals. The program is available in two versions at http://www.ualberta.ca/~augermth/piinup.html. One version requires MATLAB, but the other is a standalone version for Windows.

Our optimization of the routine and tests of accuracy used a catalogue of photographs taken in Koluktoo Bay, Nunavut, Canada (72˚02′ N, 80˚40′ W) in 2006 and 2007. The photographs were taken from land with a Nikon D70s (500 mm autofocus lens) and a Canon EOS 20D (400 mm autofocus lens). To select only photographs suitable for identification, we assigned a quality value (Q1 to Q5) to each narwhal in each photograph on the basis of five criteria: size, orientation, focus, exposure, and proportion of dorsal ridge visible in the photograph (Auger-Méthé, 2008; Auger-Méthé et al., 2010). Q5 represented images of ideal quality. For consistency with other studies (e.g., Ottensmeyer and Whitehead, 2003; Coakes and Whitehead, 2004), only individuals with a minimum of three notches and a quality rating of Q3 or above were considered suitable for matching. Photographs were matched using nicks and notches of the dorsal ridge, and because most of the photographs were taken during a short period, the pigmentation pattern and other mark types, such as scars found on the flank of the animals, were used to confirm these matches (Auger-Méthé et al., 2010). A match that is confirmed by eye is referred to as a true match. The catalogue contained 212 individuals, of which 57 were present in more than one photograph.

To compare images, the program used a description of the dorsal ridge entered manually for each image. Viewing the image on the computer screen, a user first typed a letter code and then used the mouse to designate the positions of the ends of the ridge and a set of mark points (MPs). The MP type "deep" represented the deepest points of notches, and the MP type "up" represented their upper limits (Fig. 1). Unlike notches, which cut the dorsal ridge through its entire depth, nicks only partly indent the ridge. We did not use nicks as MPs because they appear to be less stable over time than notches (Auger-Méthé et al., 2010) and their occurrence is difficult to assess reliably (Auger-Méthé, 2008).

For each MP entered, the program recorded its type (by letter code) and its position (in pixels) in relation to the x-axis of the image (Fig. 2). Once all of the MPs were entered, the proportional distance of each MP from the anterior end of the dorsal ridge was calculated as follows:

$$distMP_i = \frac{MP_i - ant\ end}{post\ end - ant\ end} \qquad (1)$$

where MP_i was the location of the *ith* MP found on the ridge, and "ant end" and "post end" the positions of the anterior and posterior ends of the dorsal ridge. When the first step of the matching process was finished, the image of the narwhal was represented by two vectors, one containing the type information for all MPs, and the other containing their distances from the anterior end of the ridge.

Using the vectors that describe its MPs, we compared each image to be matched to every other individual in the catalogue (Fig. 3). For each of these comparisons, we calculated a similarity coefficient value based on the proportion of the MPs common to both images and on how similar the locations of these MPs were. The number of common MPs was limited by the number of MPs found on the image with the fewest MPs. Each of the MPs on the image with the fewest was compared to all of the MPs of the other image, and the pairs of MPs that matched best were selected. The similarity between two MPs was based on whether two MPs of the same type were found at a similar proportional distance along the ridge, and was calculated as follows:

$$MPcomp_{ij} = e^{-\frac{1}{2}\left(\left(distMP_i - distMP_j\right)/sd\right)^2} \qquad (2)$$

where $distMP_i$ and $distMP_j$ were the proportional distances from the anterior end of the ridge for the *ith* MP of the image to be matched and the *jth* MP of the image from the catalogue. This equation was based on the probability density function of a normal distribution, and the standard deviation (SD) described the amount of error allowed in the assessment of distance. $MPcomp_{ij}$ values varied from 0 to 1, with 1 indicating that the two MPs were at exactly the same proportional distance from the anterior end of the ridge. Unlike Whitehead (1990), we did not use a threshold value to eliminate MPs that were too far apart. The $MPcomp_{ij}$ values gradually decreased with increasing distance between MPs and approached zero when the ratio between the distance and the SD was large. To calculate the overall similarity coefficient value between two images, we took the highest value selected for each common MP, summed those values, and divided the total by the average number of MPs for the two images. This averaging penalized comparisons between narwhals that differed in their number of MPs.

The principal challenge in programming narwhal photo-identification was that the ends of the dorsal ridge were hard to locate. Since we used the distance of MPs from the ends of the ridge to compare the images, errors in locating the ends could also produce matching errors. To decrease the chance of such errors, we shifted each end of each dorsal ridge to two new positions, away from the other end (lengthening the ridge), and toward the other end (shortening the ridge), and calculated the distance from each MP to the anterior end in relation to all nine ridge lengths. We repeated this process independently for each of the images compared and calculated similarity coefficient values for all combinations of ridge lengths. The highest coefficient value was selected to represent the similarity of each pair of images.

To investigate another approach to reducing errors related to ridge end location, we created a second version of the program in which the ends of the dorsal ridge were not used. Instead, the first and last MPs were the basis of the proportional distance of the MPs, and thus equation 1 was changed to:

$$distMP_i = \frac{MP_i - MP_{ant}}{MP_{post} - MP_{ant}} \qquad (3)$$

where MP_{ant} was the first MP from the anterior end of the narwhal and MP_{post} was the last MP (closest to the posterior end). The rest of the program used analogous equations and processes as described above, including the calculation of multiple sets of proportional distances for each image. As MPs can be missed by the user, the shifts in the proportional distances for this version were created to account for possible errors in the assessment of the first or last MP.

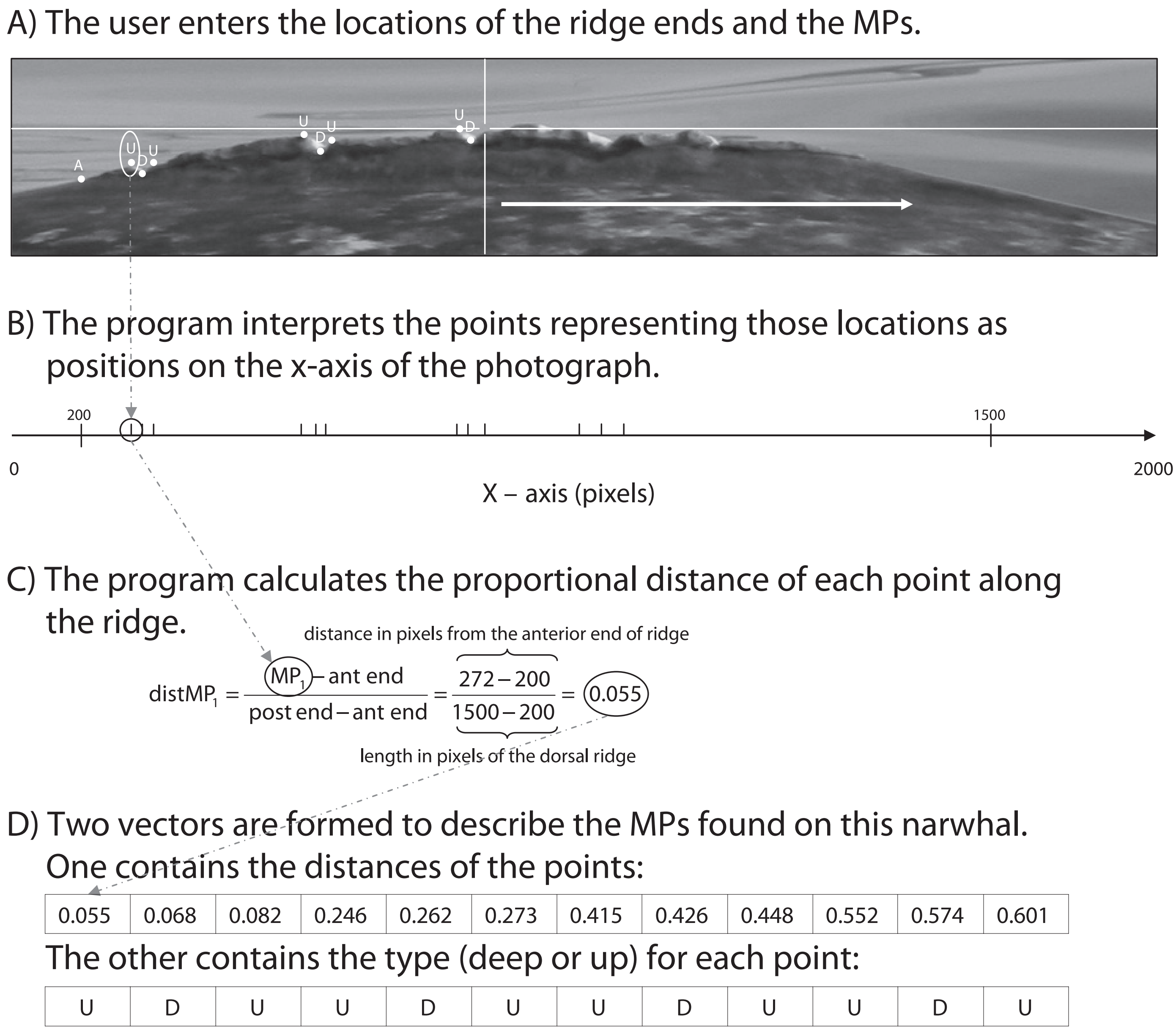

FIG. 2. Description of the first step of the matching program when using the ends of the ridge to calculate the proportional distances of MPs. The process is similar for the "no ends" version except that the first and last MPs, rather than the ends, are used to calculate the proportional distances of the MPs. A = anterior end of the ridge, U = up MPs, and D = deep MPs.

The computer program used the similarity coefficient to rank the individuals in the catalogue in decreasing order of similarity to the input image. However, it was up to the user to confirm whether the image to be matched corresponded to one of the individuals in the catalogue or whether it was a new individual. The user could choose to consider only individuals with a similarity coefficient higher or equal to a threshold value described below as potential matches. As the marks used (notches in the dorsal ridge) are visible from either side of the animal, the two sides of a narwhal could be matched to one another.

Consistency of MP Assessment

To test the consistency in the assignment of MPs to images of the same individual, we selected a sample of 80 images, representing 40 individuals. We selected only individuals that had multiple photographs of the same side taken within a year and selected a sample representing different possible pairs of quality comparison. One of the authors (M. Auger-Méthé) assigned MPs to all images in this sample. This was a visual verification of whether the user would assign the same MPs to two images of the same individual; the computer program was not used. The positions of the MPs on the ridge were recorded by placing a dot on the image using commercial imaging software, and the MP types were noted. The images were processed in

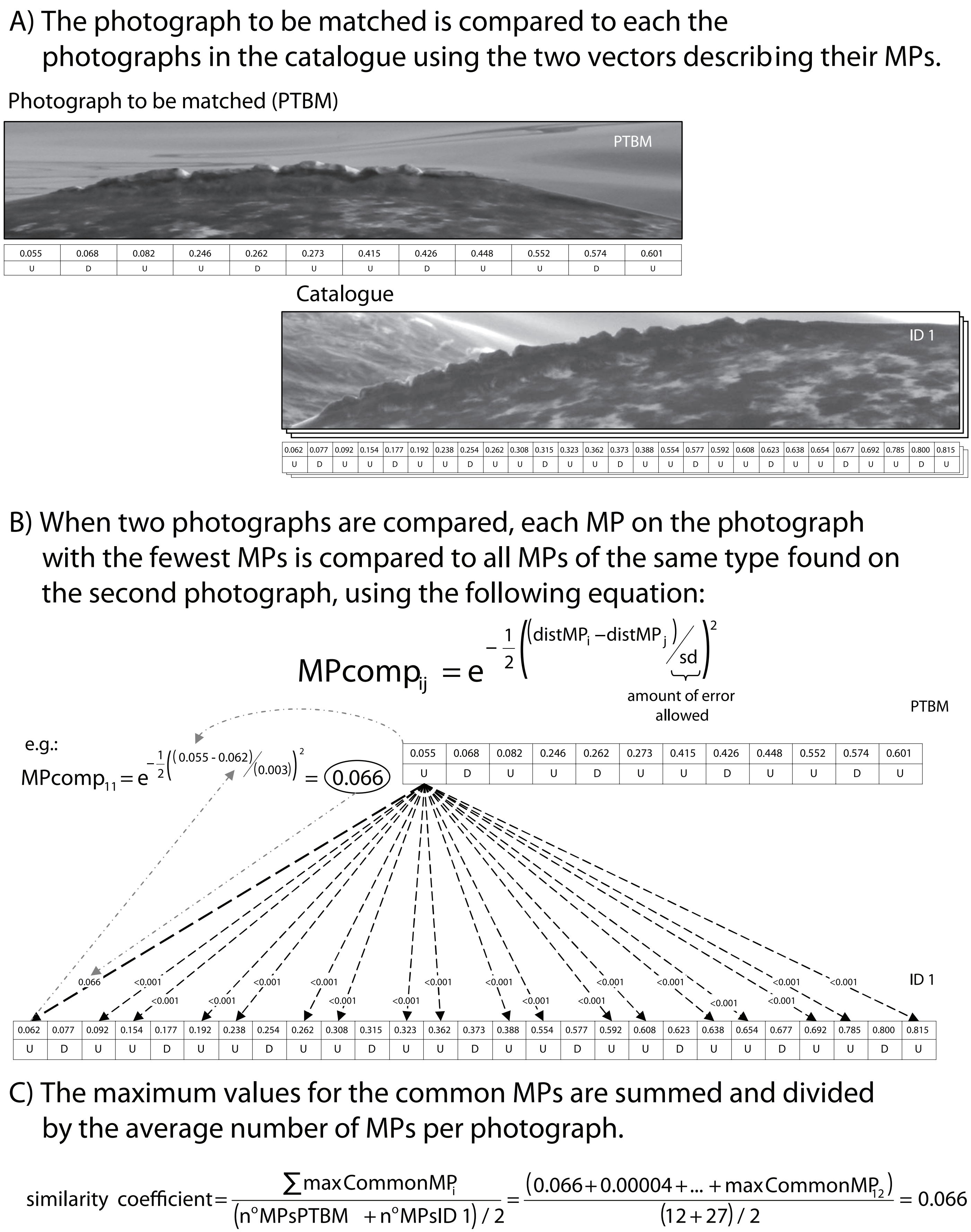

FIG. 3. Description of the second step of the matching program. U = up MPs, D = deep MPs.

random order. The images of the same individual were then compared to one another and the number of MPs the two images had in common and the number visible in only one of the images were noted. As nicks and notches are differentiated by how deeply they indent the ridge (Auger-Méthé et al., 2010), we also counted the number of missing MPs that resulted from confusion between nicks and notches. We excluded the MPs in an area that was not visible in one image of the pair (e.g., MPs on a section of the ridge that was underwater in one of the images).

For each pair of images, we calculated the proportion of missing MPs of a given type. We averaged the proportion of missing MPs across the sample pairs to calculate the error rates. To investigate whether photographic quality affected the visibility and classification of MPs, we used a Kruskal-Wallis test to compare the error rates for missing MPs across pairs of different image quality. As only two pairs of images were available, Q3 to Q5 comparisons were excluded for this test.

Consistency in Locating the Ends of the Dorsal Ridge

To assess how consistent a user was in locating the ends of the dorsal ridge, one of the authors (M. Auger-Méthé) entered the locations of these features for 30 pairs of images. Because the ends of the dorsal ridge served as the scaling landmarks of a photograph, it was difficult to estimate the discrepancy in locating these features by using different images. Thus we assessed consistency in locating the ends of the dorsal ridge by comparing two copies of the same image. To decrease the chance of remembering where we had located the ends of the dorsal ridge on the first copy, we processed the second copy a day later. Using the ends of the dorsal ridge entered for the first image of a pair as scaling landmarks, we estimated the proportional size of the discrepancy between corresponding ends.

Between-User Consistency

To assess how consistent different users are in locating the features of the dorsal ridge, two different users (M. Auger-Méthé and M. Marcoux) entered the location of those features for the same set of 30 images. By comparing how the two users assigned the features of the same image, we estimated the error rates for missing MPs and the average size of the discrepancy between the corresponding ends.

Optimization of the Program

Using the same photographic sample of 80 images as in the section *Consistency of MP Assessment*, we compared the accuracy with which three different versions of the program calculated the similarity coefficient value of two images. The first version (referred to as the "with ends" program) used the proportional distance of MPs from the ends of the ridge. The second version (the "no ends" program) used the proportional distance of MPs from the first and last MPs of the ridge, without using the ends of the ridge. The third version (the "hybrid" program) calculated both with and without the ends and selected the highest value as the similarity coefficient for the comparison.

To investigate which of these versions performed best at matching, we optimized the value of their parameters (SD in equation 2 and the proportional distances by which the ends of the ridge are shifted) and compared the optimized versions of the programs to one another. To do so, we divided the photographic sample into two sets, each containing one image of each of 40 individuals. The set containing the images to be matched was compared against all images from the set considered as the catalogue. We calculated the rate of ranking the true match among the first 10 narwhals. To optimize the parameters, we reiterated this matching process, each time altering the value of one parameter. The parameter values were chosen to optimize the rate of ranking the true match highly. This rate was also used to compare the different versions. Although the optimization of the parameters was done separately for each version, the MPs were entered only once. This procedure ensured that differences between the versions were not the results of differences in the entry of MPs, but true differences in the performances of the three versions of the matching program.

Accuracy of the Final Routine

The version of the routine found to be the best in the previous section was tested for its accuracy using a different photographic sample. We selected three sets of images: a new catalogue set, with one image for each of 40 individuals; a new match set, containing a different image of each of the same 40 individuals; and a no-match set, containing images of 30 individuals not found in the catalogue set. We selected all individuals at random from the catalogue of 212 narwhals. Since only 57 of those individuals were found in multiple photographs, about half of the images used in the match and catalogue sets of this section had been used previously to optimize the program. Thus, the accuracy calculated was more representative of the upper limits of the program's capacity than of its true accuracy. We did not limit this sample to photographs of the same side, and we included matches of narwhals photographed a year apart. We re-entered the MPs of all the images. We compared the individuals from the match and no-match sets to all of the individuals in the catalogue set. For the match set, we noted the rank and the similarity coefficient of true matches. In addition, we defined a threshold value for potential matches as the lowest similarity coefficient value noted for a true match. For each individual of the no-match set, we counted the number of comparisons against the individuals in the catalogue set that resulted in a similarity coefficient value higher than or equal to the threshold.

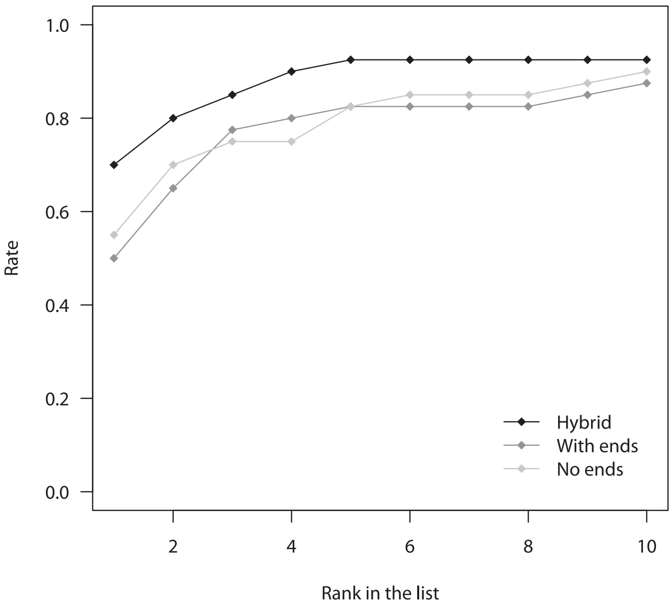

FIG. 4. Differences in matching capacity between the three optimized versions of the program. The rate corresponds to the proportion of the 40 individuals of the match set for which the true match was ranked at a value lower than or equal to each value presented on the x-axis.

Effect of Between-User Inconsistency on Matching Accuracy

To assess how robust the program was to inconsistency between users, two authors (M. Auger-Méthé. and M. Marcoux) used the computer program to extract the information from the same set of 30 images. To be able to associate these results to the between-user MP error rates and between-user discrepancies in locating the ends, we used the same 30-image set and the same placement of dorsal ridge features used to determine between-user consistency. The information entered by the first user was considered the catalogue set, and that entered by the second user, the match set. The information for each image of the match set was compared to the catalogue set. We used the final version of the computer program to calculate the similarity coefficient, and we noted the rank of the true match.

Speed

Finally, as the goal of developing this program was to reduce the time spent matching, we compared the estimated speed of the matching process with and without the program. We measured the time spent to enter the MPs of 10 individuals and the time the program took to compare these to each individual in a catalogue of 40 individuals. We also measured the time required to compare two images visually and decide whether or not they represented the same individual. Comparing two images is part of the matching process whether the program is used or not, and this last measure also served to estimate the time for matching without the program.

RESULTS

Compared to the other program versions, the hybrid had the highest rate of ranking the true match within the first 10 potential matches (0.925) (Fig. 4). Although the two other versions did not perform as well, their rates of ranking a true match within the first 10 (0.875, 0.9) were still much higher than expected by chance (0.25 for 10 out of the 40 individuals of the catalogue set). As the hybrid version consistently surpassed the others in selecting the true match within the catalogue set, it was chosen as the final version of the program. When tested with a new sample, the hybrid version ranked the true match in the first 10 potential matches at a rate of 0.90. On average, the rank of the true match was 3.75 (SE 0.79) out of 40, and thus the user on average would need to visually verify only 9% of individuals of the catalogue before finding the true match. The average value for the similarity coefficient of true matches was 0.483 (SE 0.018), and the lowest value was 0.286. If this lowest value was used as the threshold, an average of 58% of the individuals from the catalogue would be considered as potential matches for the individuals of the no-match set.

The individuals sampled to estimate the consistency in MP assessment had on average 6 deep and 12 up common MPs. The error rates in MP assessment were 0.17 (SE 0.02) missed marks per pair of images for deep and 0.15 (SE 0.02) for up. Most of the MPs missed, 94% of the deep MPs and 96% of the up MPs, were the result of confounding nicks and notches. Quality of the images (Q-value) had little effect on the up error rates (Error rates: Q3-Q3 = 0.106, Q3-Q4 = 0.188, Q4-Q4 = 0.208, Q4-Q5 = 0.078, Q5-Q5 = 0.172; Kruskal-Wallis test statistic: 6.079, df = 4, $p = 0.193$) or on the deep error rates (Error rates: Q3-Q3 = 0.103, Q3-Q4 = 0.227, Q4-Q4 = 0.250, Q4-Q5 = 0.101, Q5-Q5 = 0.178; Kruskal-Wallis test statistic: 6.824, df = 4, $p = 0.146$). The optimized distance by which the ends are shifted in the final version was 0.04 of the length of the ridge. The optimized amount of error allowed in the assessment of distance (SD in equation 2) was 0.003 of ridge length for the up MPs and 0.002 of ridge length for the deep MPs. The average distance between the locations given to the same end of the dorsal ridge in copies of the same image was 0.013 (SE 0.003) of the size of the ridge for the anterior end and 0.020 (SE 0.005) for the posterior end.

The 30 images used to estimate the consistency between users in assigning the features of the dorsal ridge had on average 4 deep and 7 up common MPs. The between-user MP error rates were 0.15 (SE 0.03) missed marks per paired copies of the same image for both deep and up MPs. The average between-user discrepancy in locating the end of the dorsal ridge was 0.051 (SE 0.014) of the size of the ridge for the anterior end and 0.058 (SE 0.011) for the posterior end. When two different users matched the same images with the computer program, the true match was ranked on average 4.60 (SE 1.23) out of 30, and its average similarity coefficient was 0.530 (SE 0.026).

When using the hybrid version, it took the user 74 sec. on average to enter the MPs, and it took the computer 0.1 sec. to compare an image to one of the individuals in the catalogue. It took on average 8 sec. to visually compare an image to one of the catalogue images and decide whether the individuals are the same or not. Using these values and the average value of 9% for the percentage of individuals that needs to be viewed before finding the true match in the catalogue, we calculated that with the computer program it would take an average of 107 sec. to find a match in a catalogue of 40 individuals and 484 sec. in a catalogue of 500 individuals. But without the computer program, since on average 50% of the catalogue would need to be viewed, the comparable figures would be 160 sec. and 2000 sec.

Adding a new individual to the catalogue required more time. We calculated that when the threshold value (0.286) is used, only 58% of the catalogue needs to be viewed before confirming that there is no match in the catalogue. Thus, it would require 264 sec. to decide that there was no match in a catalogue of 40 individuals, and 2444 sec. in a catalogue of 500 individuals. In comparison, when the program is not used, all the individuals of the catalogue would need to be viewed before deciding that there is no match in that catalogue. Thus, it would require 320 sec. to add a new individual for a catalogue of 40 individuals and 4000 sec. for a catalogue of 500 individuals. If the threshold value is not used, then using the program does not accelerate the matching process when adding a new individual.

DISCUSSION

We have developed the first photo-identification matching program for narwhals. This program differs from those developed for other species because its algorithm is designed for a nearly straight body feature with hard-to-locate scaling landmarks. The final version of the program calculated the similarity coefficient using both the proportional distance from the ends of the ridge and the proportional distance from the first and last MPs. The hybrid version of the program was selected because it consistently surpassed the others in ranking the true match among the first 10 individuals in the list. The absolute accuracy of our program, measured as the proportion of true matches that are ranked first, was 52.5%. This was comparable to the absolute accuracy of programs developed for other species, which ranges between 32% and 72% (Whitehead, 1990; van Tienhoven et al., 2007; Gamble et al., 2008). The relative accuracy, which was shown by Gamble et al. (2008) to be constant with increasing catalogue size, was measured as the proportion of true matches that are ranked within the first 10% of the catalogue. Our program's relative accuracy was 77.5%, which is within the 50%–80% range achieved by other matching programs (Stewman et al., 2006; Finerty et al., 2007).

Identifying a new individual (an individual with no match in the catalogue) is the most time-consuming task (Defran et al., 1990), as the image generally has to be compared to all of the individuals in the catalogue. Therefore, finding a threshold value of the similarity coefficient below which the user does not need to look for potential matches is also an important measure of accuracy. Most programs do not employ a threshold value that limits the number of individuals considered to be potential matches. However, using such thresholds can be effective in reducing the time spent matching. For example, Hiby and Lovell (1990) found that true matches for grey seal (*Halichoerus grypus*) had a similarity coefficient greater than 0.5 and that 98% of the comparisons between two different individuals had a similarity coefficient lower than 0.5. Such a clear threshold could not be found for our routine. The lowest similarity coefficient for a true match was 0.286, and on average, 58% of the comparisons between two different individuals had a similarity coefficient equal to or greater than 0.286. Although this threshold does not completely distinguish true matches from incorrect matches, its use still decreases the number of individuals that need to be verified visually.

Part of the limitation on program accuracy can be explained by inconsistencies in MP assessment and in locating the ends of the dorsal ridge. MP assessment errors were common, and most of these errors resulted from assessing nicks, which are shallower indents in the ridge, as notches. As expected, locating the ends of the dorsal ridge was difficult. Locating the posterior end of the ridge was particularly difficult because, unlike the abruptly rising anterior end, the posterior end gradually merged with the back of the narwhal. A combination of errors in placing MPs and in locating the ends of the ridge explained why some true matches had low similarity values and ranked low in the list of potential matches.

Discrepancies in locating the features of the dorsal ridge will likely increase as the number of program users increases. The discrepancies in locating the ends of the dorsal ridge were greater for between-user assignment than for repeated assignments by the same user. In contrast, the between-user MP error rates were similar to the same-user MP error rates. Note, however, that the between-user MP error rates were estimated using copies of the same images, while the same-user MP error rates were estimated using different images of the same individual. Thus the between-user MP error rates might be underestimated compared to the same-user MP error rate. Because of the between-user discrepancies, the average rank of the true match was lower when different users matched the same image set than when the same user matched different images of the same individual. It is thus very important that the guidelines on identifying MPs be followed. It is possible that the development of a method that automatically retrieves information from the image (see Hiby and Lovell, 1990; Hillman et al., 2003; Arzoumanian et al., 2005; Gamble et al., 2008) would reduce such MP assessment errors, and this possibility could be explored in the future.

Identification Errors

Photo-identification has two inherent types of identification errors: (1) matching two different individuals as the same one (false positive) and (2) considering one individual as two (false negative). Both are generally the consequence of using non-distinct marks, photographs of poor quality, or marks that change with time (Carlson et al., 1990; Agler, 1992) and can lead to bias in studies using photo-identification (Hammond, 1986; Gunnlaugsson and Sigurjónsson, 1990; Stevick et al., 2001). As the user is required to visually confirm matches, the rate of false-positive errors should not be affected by using the program, compared to matching without it. However, the rate of false-negative errors could be increased with the use of a threshold value that limits the number of individuals considered as potential matches. One method to correct for the bias resulting from such false-negative errors is to include the probability of identification error as a parameter in the capture-recapture model of interest (e.g., Yoshizaki et al., 2009). Using models that incorporate the probability of making false-negative errors allows one to use the threshold without misgivings regarding the possible bias such identification errors could introduce in the study. Users who prefer to use simpler models can easily avoid the increase in false-negative errors resulting from using the computer program by ignoring the threshold and considering all images of the catalogue as potential matches. However, if the threshold is ignored, using the computer program will not accelerate the matching process when the image to be matched is of a new individual.

A decrease in distinctiveness of the features used in matching has been shown to reduce the ability of matching programs to rank true matches first (Beekmans et al., 2005; Finerty et al., 2007). One factor that can increase the distinctiveness of an individual is an increase in the number of reference points that are used to compare two images. Other studies have limited their photographic sample to individuals with a minimum of three references points (Ottensmeyer and Whitehead, 2003; Auger-Méthé and Whitehead, 2007). Although up MPs help describe the features of notches, deep MPs are the main features describing notches, thus making three deep MPs a suitable restriction for the distinctiveness of individual narwhals. Note that since calves and many juveniles do not have notches (Auger-Méthé et al., 2010), this criterion will exclude young narwhals from the catalogue.

A decrease in accuracy of other programs was also associated with a decrease in photographic quality (Whitehead, 1990; Beekmans et al., 2005; Finerty et al., 2007), and a change in angle was found to be the main factor affecting the accuracy of the program (Whitehead, 1990; Kelly, 2001). Although the lack of differences in the error rates associated with images of different quality indicates that the marks are equally visible across images of different quality, the effect the angle might have on the program was not tested specifically. Although using only pictures rated Q3 or above limits the possible amount of angle the ridge can have in a picture, subtle differences in angle could still potentially affect the proportional distance of MPs.

Although notches appear to be relatively stable over time (Auger-Méthé et al., 2010), their long-term stability has not been properly assessed. As changes in notches could affect both the accuracy of the computer program and the number of misidentifications, the rate of change of notches and the rate of misidentification should be formally investigated. While incorporating an estimate of the rate of misidentification can significantly reduce biases in models, such as population estimate models (e.g., Stevick et al., 2001), one can incorporate misidentification due to changes in marks in models without knowing the rate of identification error (e.g., Yoshizaki et al., 2009).

Speed

Although high accuracy is an important attribute of a matching program, such a program is useful only if it decreases the time spent in matching. The computer-assisted matching process was estimated to be 1.5 times as fast at finding a match in a catalogue of 40 individuals, and 4.1 times as fast for a catalogue of 500 individuals. Although the difference in speed was not as large when assessing new individuals, with the computer program the process was still about 1.2–1.6 times as fast as visual matching. The computer program accelerated the matching process of a new individual only if the threshold value described above was used. The time required for a program to match individuals is dependent on the processing power of the computer on which it is used. Most computers available on the market should have higher processing power than the one used to estimate the speed of the program, a Compaq Presario 900. Therefore, we can confidently say that the program will accelerate the matching process, especially with a large catalogue.

CONCLUSION

Conservation agencies have emphasized that data deficiency and uncertainties in abundance estimates and trends of narwhals are causes for concern (COSEWIC, 2004; Jefferson et al., 2009). Although current research methods used to study narwhals have significantly increased our knowledge of the species, they are generally expensive or invasive. With recent recognition that invasive methods are condemned by some Inuit communities (George, 2006; Siku News, 2007) and can have negative impacts on the individuals studied (e.g., Saraux et al., 2011), there is an increased need to facilitate the use of non-invasive methods to study northern species. Although photo-identification is a relatively inexpensive and non-invasive method to study varied aspects of the narwhal's ecology, matching individuals is a time-consuming task. To alleviate the main drawback of photo-identification, we have developed a computer program that accelerates the photo-identification of narwhals. Beluga whales (*Delphinapterus leucas*), like narwhals, have dorsal ridges that are sometimes notched and possess hard-to-locate ends. Thus our program could be applied to this species. Our matching program has a graphical user interface that is simple to use and thus it could facilitate the integration of photo-identification into a community-based monitoring program, which could allow northern communities to be more involved in managing a cetacean population that they exploit.

ACKNOWLEDGEMENTS

We thank the Canadian Marine Environment Protection Society, the Canadian Whale Institute, the Canadian Wildlife Federation, the Department of Fisheries and Oceans Canada, the Eben Hopson Fellowship, Catalina Gomez-Salazar, Murray Humphries, Namen Inuarak, Abraham Kublu, Jean-Sébastien Lauzon-Guay, the Lorraine Allison Scholarship, the Mittimatalik Hunters and Trappers Organization, the Natural Sciences and Engineering Research Council, Nikon Canada, the Northern Scientific Training Program of Indian and Northern Affairs Canada, Polar Sea Adventures, Dave Reid, Cindy Staicer, Luke Suqslak, and the World Wildlife Fund Canada for their support or assistance. We also thank Philip Hammond, Jeff Hutchings, Devin Lyons, and Ian McLaren for their useful comments on previous versions of the manuscript. Our work made use of the resources of the Academic Information and Communication Technologies of the University of Alberta.

REFERENCES

Agler, B.A. 1992. Testing the reliability of photographic identification of individual fin whales (*Balaenoptera physalus).* Report of the International Whaling Commission 42:731–737.

Arzoumanian, Z., Holmberg, J., and Norman, B. 2005. An astronomical pattern-matching algorithm for computer-aided identification of whale sharks *Rhincodon typus*. Journal of Applied Ecology 42(6):999–1011.

Auger-Méthé, M. 2008. Photo-identification of narwhals. MSc thesis, Dalhousie University, Halifax, Nova Scotia.

Auger-Méthé, M., and Whitehead, H. 2007. The use of natural markings in studies of long-finned pilot whales (*Globicephala melas*). Marine Mammal Science 23(1):77–93.

Auger-Méthé, M., Marcoux, M., and Whitehead, H. 2010. Nicks and notches of the dorsal ridge: Promising mark types for the photo-identification of narwhals. Marine Mammal Science 26(3):663–678.

Beekmans, B.W.P.M., Whitehead, H., Huele, R., Steiner, L. and Steenbeek, A.G. 2005. Comparison of two computer-assisted photo-identification methods applied to sperm whales (*Physeter macrocephalus*). Aquatic Mammals 31(2):243–247.

Carlson, C.A., Mayo, C.A., and Whitehead, H. 1990. Changes in the ventral fluke pattern of the humpback whale (*Megaptera novaeangliae*), and its effect on matching; evaluation of its significance to photo-identification research. In: Hammond, P.S., Mizroch, S.A., and Donovan, G.P., eds. Use of photo-identification and other techniques to estimate population parameters. Reports of the International Whaling Commission, Special Issue 12. 105–111.

Coakes, A.K., and Whitehead, H. 2004. Social structure and mating system of sperm whales off northern Chile. Canadian Journal of Zoology 82(8):1360–1369.

COSEWIC (Committee on the Status of Endangered Wildlife in Canada). 2004. COSEWIC assessment and update status report on the narwhal *Monodon monoceros* in Canada. Ottawa: COSEWIC.

Defran, R.H., Shultz, G.M., and Weller, D.W. 1990. A technique for the photographic identification and cataloging of dorsal fins of the bottlenose dolphin (*Tursiops truncatus*). In: Hammond, P.S., Mizroch, S.A., and Donovan, G.P., eds. Use of photo-identification and other techniques to estimate population parameters. Reports of the International Whaling Commission, Special Issue 12. 53–55.

Finerty, S.E., Hillman, G.R., and Davis, R.W. 2007. Computer-matching of sea otter (*Enhydra lutris*) nose scars: A new method for tracking individual otters. Aquatic Mammals 33(3):349–358.

Gamble, L., Ravela, S., and McGarigal, K. 2008. Multi-scale features for identifying individuals in large biological databases: An application of pattern recognition technology to the marbled salamander *Ambystoma opacum*. Journal of Applied Ecology 45(1):170–180.

George, J. 2006. NWMB picky about support for scientists. *Nunatsiaq News*, April 7. http://www.nunatsiaqonline.ca/stories/article/nwmb_picky_about_support_for_scientists.

Gunnlaugsson, T., and Sigurjónsson, J. 1990. A note on the problem of false positives in the use of natural marking data for abundance estimation. In: Hammond, P.S., Mizroch, S.A., and Donovan, G.P., eds. Use of photo-identification and other techniques to estimate population parameters. Reports of the International Whaling Commission, Special Issue 12. 143–145.

Hammond, P.S. 1986. Estimating the size of naturally marked whale populations using capture-recapture techniques. In: Donovan, G.P., ed. Behaviour of whales in relation to management. Reports of the International Whaling Commission, Special Issue 8. 253–282.

Hiby, L., and Lovell, P. 1990. Computer aided matching of natural markings: A prototype system for grey seals. In: Hammond, P.S., Mizroch, S.A., and Donovan, G.P., eds. Use of photo-identification and other techniques to estimate population parameters. Reports of the International Whaling Commission, Special Issue 12. 57–61.

Hillman, G.R., Würsig, B., Gailey, G.A., Kehtarnavaz, N., Drobyshevsky, A., Araabi, B.N., Tagare, H.D., and Weller, D.W. 2003. Computer-assisted photo-identification of individual marine vertebrates: A multi-species system. Aquatic Mammals 29(1):117–123.

Jefferson, T.A., Karczmarski, L., Laidre, K., O'Corry-Crowe, G., Reeves, R.R., Rojas-Bracho, L., Secchi, E.R., et al. 2009. *Monodon monoceros*. In: IUCN Red List of Threatened

Species. Version 2010.2. http://www.iucnredlist.org/apps/redlist/details/13704/0.

Kelly, M.J. 2001. Computer-aided photograph matching in studies using individual identification: An example from Serengeti cheetahs. Journal of Mammalogy 82(2):440 – 449.

Marcoux, M., Auger-Méthé, M., and Humphries, M.M. 2009. Encounter frequencies and grouping patterns of narwhals in Koluktoo Bay, Baffin Island. Polar Biology 32(12):1705 – 1716, doi:10.1007/s00300-009-0670-x.

Mathworks. 2007. Matlab: The language of technical computing. Natick, Massachusetts: MathWorks Inc.

Ottensmeyer, C.A., and Whitehead, H. 2003. Behavioural evidence for social units in long-finned pilot whales. Canadian Journal of Zoology 81(8):1327 –1338.

Saraux, C., Le Bohec, C., Durant, J.M., Viblanc, V.A., Gauthier-Clerc, M., Beaune, D., Park, Y.-H., Yoccoz, N.G., Stenseth, N.C., and Le Maho, Y. 2011. Reliability of flipper-banded penguins as indicators of climate change. Nature 469:203 – 206.

Siku News. 2007. Inuit condemn invasive research practices. *Siku News*, December 7. http://www.sikunews.com/News/Canada-Nunavut/4125.

Stevick, P.T., Palsbøll, P.J., Smith, T.D., Bravington, M.V., and Hammond, P.S. 2001. Errors in identification using natural markings: Rates, sources, and effects on capture-recapture estimates of abundance. Canadian Journal of Fisheries and Aquatic Sciences 58(9):1861 – 1870.

Stewman, J., Debure, K., Hale, S., and Russell, A. 2006. Iterative 3-D pose correction and content-based image retrieval for dorsal fin recognition. Lecture Notes in Computer Science 4141:648 – 660.

Van Tienhoven, A.M., den Hartog, J.E., Reijns, R.A., and Peddemors, V.M. 2007. A computer-aided program for pattern-matching of natural marks on the spotted raggedtooth shark *Carcharias taurus*. Journal of Applied Ecology 44(2):273 – 280.

Whitehead, H. 1990. Computer assisted individual identification of sperm whale flukes. In: Hammond, P.S., Mizroch, S.A., and Donovan, G.P., eds. Use of photo-identification and other techniques to estimate population parameters. Reports of the International Whaling Commission, Special Issue 12. 71 – 77.

Yoshizaki, J., Pollock, K.H., Brownie, C., and Webster, R.A. 2009. Modeling misidentification errors in capture–recapture studies using photographic identification of evolving marks. Ecology 90(1):3 – 9.