Journal of FISH BIOLOGY

fsbi An International Society for Fish Biology

Journal of Fish Biology (2016)
doi:10.1111/jfb.13039, available online at wileyonlinelibrary.com

Automated otolith image classification with multiple views: an evaluation on Sciaenidae

J. Y. Wong*, C. Chu*, V. C. Chong*†‡, S. K. Dhillon* and K. H. Loh†

*Institute of Biological Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia and †Institute of Ocean and Earth Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia

(Received 30 October 2014, Accepted 13 April 2016)

Combined multiple 2D views (proximal, anterior and ventral aspects) of the sagittal otolith are proposed here as a method to capture shape information for fish classification. Classification performance of single view compared with combined 2D views show improved classification accuracy of the latter, for nine species of Sciaenidae. The effects of shape description methods (shape indices, Procrustes analysis and elliptical Fourier analysis) on classification performance were evaluated. Procrustes analysis and elliptical Fourier analysis perform better than shape indices when single view is considered, but all perform equally well with combined views. A generic content-based image retrieval (CBIR) system that ranks dissimilarity (Procrustes distance) of otolith images was built to search query images without the need for detailed information of side (left or right), aspect (proximal or distal) and direction (positive or negative) of the otolith. Methods for the development of this automated classification system are discussed.

Key words: automated classification; geometric morphometrics; Malaysia; otoliths; shape analysis.

INTRODUCTION

Species in the family Sciaenidae, commonly called croakers or drums, are found in both marine and fresh waters of the tropical and temperate zones. Except for a few species in coral reefs, most are euryhaline inhabiting turbid coastal waters, bays, estuaries and rivers (Sasaki, 2001). Currently, the family includes 290 species in 66 genera worldwide (Eschmeyer *et al.*, 2010) and in Malaysia alone, 25 species in 12 genera were reported (Chong *et al.*, 2010). Positive identification of sciaenid specimens to genus and species requires the examination of swimbladders and sagittal otoliths. The swimbladder is perhaps the most useful for identification, being characterized as carrot or hammer-shaped and it may bear lateral branching appendages depending on genus and species (Sasaki, 1989). The form of the sulcus acusticus on the proximal surface of the otolith may be a distinguishing feature among species of fishes (Tuset *et al.*, 2008). It is generally less useful in the Sciaenidae, however, except for the genera *Argyrosomus* L. 1758, *Otolithes* Oken 1817 and *Pterotolithus* Fowler 1933 (Sasaki, 2001). The posterior crenulations or the marginal domes of the otoliths may be distinctive between

‡Author to whom correspondence should be addressed. Tel.: +60 379674220; email: chong@um.edu.my

species, but their positional descriptions in a taxonomic key are tenuous and rarely adopted in conventional fish keys.

Since the 1990s, morphometric analysis has increasingly been used to discriminate the species, population, sex or age of fishes (Corti & Crosetti, 1996; Cavalcanti, *et al.*, 1999; Herler *et al.*, 2010). Accordingly, automated systems based on digital imaging have rapidly developed to increase the efficiency and usage of such analyses in fish identification, whether by fishery scientists or non-specialists (Hu *et al.*, 2012). Semi-automated systems, however, incorporating conventional meristic and morphometric characters, have been proposed as opposed to completely replacing conventional characters with digital imaging (Guisande *et al.*, 2010). Since otolith structure is much simpler than the whole fish morphology, otolith images have been routinely used to develop automated classification systems (Parisi-Baradad *et al.*, 2010).

Currently, a few online databases of otolith digital images exist, such as AFORO (Anàlisi de Formes d'otòlits; www.cmima.csic.es), which contains species of fishes from the Mediterranean and Antarctic waters (Parisi-Baradad *et al.*, 2010) and the *Otolith Atlas of Taiwan*, which mostly covers species from and near Taiwan (Lin & Chang, 2013). The AFORO database integrates a content-based image retrieval (CBIR) system in which the query image is automatically searched against similar otolith images found in the database. Such an automated system helps users to search for otolith images and identify them even without good familiarity with the otolith morphology. In the present study, a CBIR system is proposed based on Procrustes analysis (Goodall, 1991), with which biologists are more familiar than other methods such as curvature scale space (CSS) and wavelet methods adopted in AFORO. The aim is to make the search easier for the non-specialist, while requiring less information about the query (otolith).

Besides incorporating the CBIR system, detailed three-dimensional scans are also available in the AFORO database. Shape information found in the 3D structure is irrefutably useful for classification, but the use of 3D scanners is costly and requires longer scanning times. One of the convenient ways to capture partial 3D information is by combining multiple 2D images (Fadda *et al.*, 1997; Chen *et al.*, 2008). In the present study, classification performance of nine species of Sciaenidae using single view and combinations of 2D views of the otolith was evaluated.

MATERIALS AND METHODS

SPECIMENS

Fish specimens from nine species in eight genera of the family Sciaenidae were used. These included goatee croaker *Dendrophysa russelii* (Cuvier 1829) ($n = 30$), Belanger's croaker *Johnius belangerii* (Cuvier 1830) ($n = 29$), sharpnose hammer croaker *Johnius borneensis* (Bleeker 1851) ($n = 15$), Caroun croaker *Johnius carouna* (Cuvier 1830) ($n = 32$), soldier croaker *Nibea soldado* (Lacépède 1802) ($n = 21$), tigertooth croaker *Otolithes ruber* (Bloch & Schneider 1801) ($n = 41$), Panna croaker *Panna microdon* (Bleeker 1849) ($n = 36$), donkey croaker *Pennahia anea* (Bloch 1793) ($n = 13$) and blotched tiger-toothed croaker *Pterotolithus maculatus* (Cuvier 1830) ($n = 6$). Species were identified according to the FAO identification key (Sasaki, 2001). A total of 223 right side otoliths were extracted (Table SI, Supporting Information).

Otoliths were viewed under a Olympus DP25FW, ×6·3 magnification stereomicroscope (Olympus; www.olympus.com) attached with a camera connected to a computer. The proximal

aspect of the right otolith, with its anterior end facing the negative direction on a Cartesian co-ordinate system and dorsal edge facing up, was photographed (standard image). A circular fluorescent lamp was used to enhance object contrast with the dark background. Otolith images were further captured at two other aspects: anterior and ventral. The anterior image had the dorsal edge facing in the negative direction, whereas the ventral image had the anterior end facing in the negative direction. The image of the proximal aspect was used for all analyses. Images of the anterior and ventral aspects were included for the purpose of comparison and to enhance classification (Fig. S1, Supporting Information).

SOFTWARE

All of the implementations described in the following sections (except evaluation of allometric effect) were performed using the package otolith 0.2.3 written for this study. The package was written to run in R (www.r-project.org) and the source codes and documentations can be downloaded from an online repository (http://github.com/jinyung/otolith/). This package imported codes from other works, mostly from morpho 2.2 in R (Schlager, 2014) and using functions developed by Claude (2008) for morphometrics and EBImage 4.6.0 (Pau *et al.*, 2010) for image processing.

IMAGE PROCESSING AND OUTLINE SAMPLING

Digital images of the otoliths were first converted into grey-scale images and thresholded. Objects not of interest (*e.g.* artefacts and dust specks) were filtered out by area. The otolith outline was traced and saved as Cartesian co-ordinates. Two outline sampling schemes, the comb method (Ponton, 2006) and the equidistant method (Lombarte *et al.*, 2010), were tested. Since the comb method could not sample from folded outlines (*e.g.* at the excisura site) efficiently (Fig. S2, Supporting Information), a modified equidistant method was used. In the modified equidistant method, the two farthest end points of the otolith were initially defined. The line that joins them represents the major principal axis of the outline. This major principal axis is approximately the biological anterior–posterior axis that bisects the otolith outline into upper and lower halves. Fifty points (referred to as semi-landmarks) were sampled along the otolith outline at equally spaced distances in the clockwise direction for each half, starting from the left end point of the major principal axis. A total of 100 Cartesian co-ordinates of semi-landmarks were sampled (Fig. 1) and these co-ordinates are referred to as configurations (Claude, 2008). The same outline sampling scheme was used for images obtained from all three different aspects (*i.e.* proximal, anterior and ventral) of the otolith. Assuming that the left–right side otolith, or the facing direction (negative–positive) of the otolith are unknown, a few possible semi-landmark configurations could result from this sampling algorithm for a pair of otoliths in an individual, as presented in Fig. 2. Given these possible configuration types, it is possible to develop a search function that considers all the variants of the query otolith image.

SHAPE DESCRIPTION

The sampled semi-landmarks were analysed using geometric morphometric methods, the generalized Procrustes analysis (GPA; Rohlf & Slice, 1990) and elliptical Fourier analysis (EFA; Rohlf & Archie, 1984). GPA and EFA were used to describe the shape of the outline using the same sets of semi-landmarks. For the GPA method, all semi-landmark configurations were first subjected to sliding to achieve better homology among the specimens. Minimization of bending energy was used as the criterion in semi-landmark sliding (Bookstein, 1997). In addition to geometric morphometrics, shape indices were also calculated for each otolith (Ponton, 2006; Tuset *et al.*, 2006). A macro was written for this purpose to run in ImageJ (Schneider *et al.*, 2012), utilizing the Particle 8 plug-in written by Landini (2008); available at www.mecourse.com/landinig/software/software.html to perform the calculations. The shape indices include Feret's aspect ratio, circularity, roundness, compactness, solidity, convexity, shape, *R*-factor, modification ratio, sphericity, rectangularity and fractal dimension. All the indices were calculated using the plug-in, except for fractal dimension (box-count method),

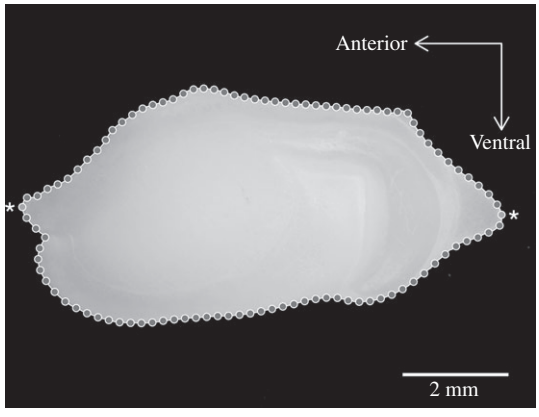

FIG. 1. Semi-landmark sampling method used on the sagittal otoliths of nine sciaenid species. Semi-landmarks were sampled on the outline of the otolith starting from the left end point of the major principal axis (*) and ending at the same end point in the clockwise direction.

which was included in base ImageJ 1.38r. Definitions of the indices can be found at the plug-in website and ImageJ website (http://rsb.info.nih.gov/ij/). All three methods are scale and rotation-invariant. While GPA inherently removes scale (here, unit size scaling was used, referred to as partial GPA) and rotation, a normalization procedure (including starting point and rotation angle) was used for EFA (referred to as normalized EFA). Computationally, implementations by Schlager (2014) and Claude (2008) were adopted for GPA and normalized EFA, respectively. Only the retained shape information was used for subsequent model training. The size measurements used for scale normalization, namely the centroid size (from GPA) and the magnitude of the semi-major axis of the first fitting ellipse (from EFA), were also calculated and used in subsequent analyses.

CLASSIFICATION MODELS AND EVALUATION OF CLASSIFICATION PERFORMANCE

Classification models were trained using either principal component scores calculated from the GPA-rotated semi-landmarks, coefficients of the harmonic series of EFA or calculated shape indices. Linear discriminant analysis (LDA) was used as the classifier for all the evaluations using equal priors for all classes. Classification models were evaluated using iterated *k*-fold cross-validation to provide a reasonably stable and low biased measure of model performance (Bouckaert, 2003). While comparison of model performance was based on overall accuracy, evaluation of prediction performance in each class (species) was based on recall, precision and specificity. Recall (also called sensitivity) measures the proportion of specimens correctly classified as a species out of all specimens from that species; precision (also called positive predictive value) measures the proportion of specimens correctly classified as a species out of all specimens classified as that species by the classifier; specificity measures the proportion of specimens correctly classified as 'not a species' out of all specimens not belonging to that species (see also Sokolova & Lapalme, 2009 for more discussion on their interpretations and extensions). Unless otherwise stated, all performance measures from cross-validation were estimated from the average of over 100 iterations of five-fold cross-validations, *i.e.* a total of 500 models. Evaluation of performance on specimen level was based on the proportion of correct predictions out of all iterations. Each specimen was predicted exactly once in each iteration for iterated *k*-fold cross-validation. The effect of reducing the number of principal components (PC) or harmonic series was usually evaluated using per cent of variation explained as the indicator of goodness of fit. Here, the overall accuracy was used as a measure since classification was the main objective.

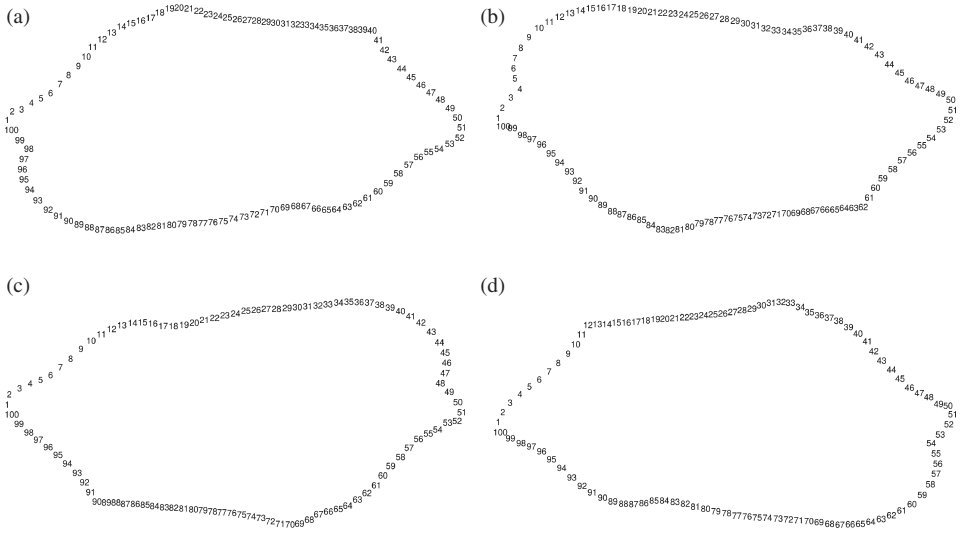

FIG. 2. Possible semi-landmark configuration types under current sampling method for otolith of *Panna microdon*. There are eight possible types: (a) proximal aspect of the right otolith with anterior end pointing to negative direction (standard image), or distal aspect of the left otolith with anterior end pointing to negative direction, (b) proximal view of the left otolith with anterior end pointing to positive direction, or distal view of the right otolith with anterior end pointing to negative direction [*i.e.* 2(a) flipped over], (c) proximal view of right otolith with anterior end pointing to positive direction, or distal view of left otolith with anterior end pointing to positive direction and (d) proximal view of the left otolith with anterior end to the positive direction, or distal view of the right otolith with anterior end pointing to positive direction [*i.e.* 2(c) flipped over].

For evaluation of dimension reduction, different numbers of principal components or harmonic series were evaluated using iterated k-fold cross-validation (30 iterations) and the results were compared to select a suitable range of PCs or harmonic numbers after reducing the number of dimensions. In an attempt to further improve the performance, an ensemble approach based on the iterated k-fold resampling method was used to aggregate the classification models (Beleites & Salzer, 2008). In brief, the predictions from the surrogate models that resulted from the iterated k-fold resampling method were aggregated to provide prediction through majority votes, *i.e*. each surrogate model was trained with only a part of the training data (4/5 if $k = 5$) to make predictions and the species with the most votes out of 500 votes (if 100 iterations of five-fold resampling was used) was reported as the final prediction. The aggregated method is the same as the bagging method (bootstrap aggregating; Breiman, 1996), but with a different resampling method (iterated k-fold sampling *vs*. bootstrapping). No further cross-validation was used for variance estimates for aggregated LDA model.

COMPARISON OF DIFFERENT OTOLITH VIEWS

Shape data from different views were combined by concatenating their data matrices (PC scores, EFA coefficients or shape indices), where each data matrix has its own chosen parameters (*e.g*. range of dimension reduction). No further parameter tuning was performed.

REMOVING UNCERTAIN PREDICTIONS

Prediction made by LDA for each specimen was based on the class that gave the maximum *a posteriori* probability estimate (Venables & Ripley, 2002). Utilizing this condition, it is possible to remove prediction of low confidence by thresholding the posterior probability. Different

threshold values of the posterior probability were evaluated and the overall accuracy and total number of accepted predictions were calculated. The overall accuracy was calculated from the true positives out of the accepted predictions after exclusion, termed the accepted accuracy (Ye *et al.*, 2011). This post prediction strategy was also used for the aggregated model, but using a different criterion. Thresholding was done based on the proportion of votes from the aggregated model, excluding votes below the threshold value.

EVALUATION OF ALLOMETRIC EFFECT

Although geometric morphometric methods can remove size differences mathematically, otoliths of different sizes still possess considerable within-species shape differences. Evaluation of allometry was performed through multivariate regression of otolith shape variations on otolith size variations. Procrustes distance was used as the measure for shape variations and centroid size was used as the measure for size variations. The possible effect of otolith allometry on classification performance was qualitatively checked using PCA bubble plots, where PCA was performed on the shape data and otolith size (centroid size) was reflected by bubble size. The effect of species on otolith shape variations was also tested using multivariate analysis of covariance (MANCOVA), using otolith size as covariate. Multivariate regression and MANCOVA were performed using geomorph 2.0.1 in R (Adams & Otárola-Castillo, 2013).

EVALUATION USING THE TEST DATA SET

In addition to cross-validation, a dedicated test set of otoliths was also used to verify the performance of the classification model derived from the training set. The test set consisted of the same three otolith images from each species (except for *J. borneensis*, with only a single image), for a total of 25 images independent of the training set.

Two variants of the test set were prepared using otolith images from the same set of fish specimens. One set consists of images of the right otolith in the standard direction (*i.e.* anterior end pointing to negative direction for proximal images and ventral images; ventral edge pointing to positive direction for anterior image; referred to as standard test set). The second set consists of images of either left or right otolith, with anterior end (for proximal and ventral images) or ventral edge (for anterior image) randomly assigned to negative or positive direction (referred to as random test set). Assessment results of the random test set were also used to determine the prediction accuracy regarding otolith view and direction by the search algorithm. All test sets were evaluated using classification models trained with the full training data set with no further tuning of parameters.

SEARCH AND PREDICTION OF UNKNOWN SPECIMENS

A search engine was developed to enable image search of new, unknown specimens against an otolith database of known species (training set). Procrustes distance was used as a distance metric to rank the similarity between the query image and the training set, *i.e.* the smaller the pair-wise distance, the higher the ranked similarity.

To enable a search without prior knowledge of side and direction of the otolith, an optional feature under the current semi-landmark sampling scheme was developed. It uses an exhaustive searching algorithm to search configurations for either left or right otolith, proximal or distal surface and negative or positive direction of anterior end (Fig. 2). The query was initially searched against the mean shape configurations of each species in the database to shorten the search time. Hence, mean shape configurations were first computed and saved in the database. For each species in the database, the range of Procrustes distance between specimens and species' mean shape was also calculated.

The search then determined whether the Procrustes distance between the query (all possible configurations) and species' mean shape fell within the range of a species. For example, if the query fell within the ranges of two species, the search will be limited to these two species only. Otherwise, the search will continue with the specimens of five closest species to reduce search time. The search finally returned to the closest otolith matches, with guesses of the side and

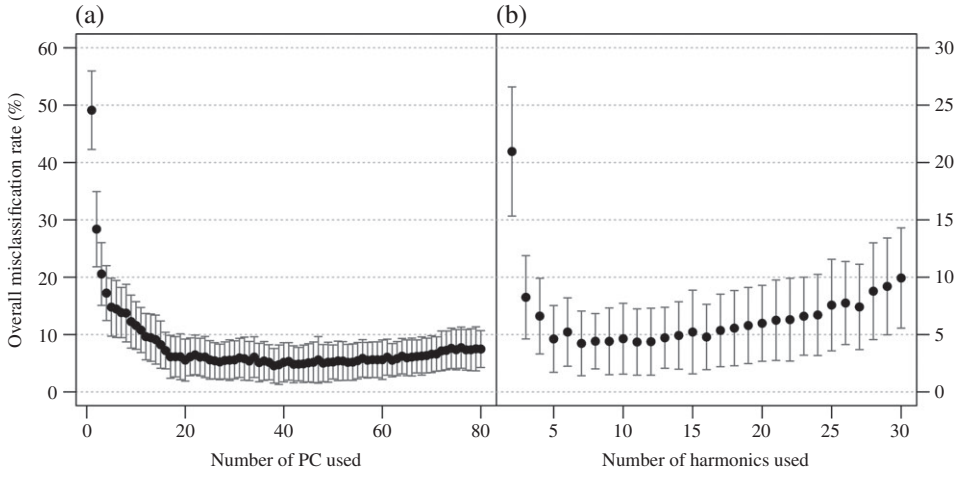

FIG. 3. Effect of dimension reduction on classification performance, expressed as changes in overall misclassification rate (mean ± S.D.). Examples shown here are the evaluations performed on (a) the generalized Procrustes analysis (GPA) method and (b) the elliptical Fourier analysis (EFA) method.

direction of the otolith. To test the search algorithm, randomly rotated and flipped images in the training set were used as queries. The test search was repeated 10 times. The search algorithm was then used as an option to predict the direction and side of the query image before subjecting it to the classification model. A flowchart of how otolith images were processed, computed, analysed and classified is shown in Fig. S3 (Supporting Information).

RESULTS

EFFECT OF DIMENSION REDUCTION

The effect of using different numbers of PC (for the GPA method) and harmonic series (for the EFA method) on the overall classification accuracy of otoliths is shown in Fig. 3. For both methods, the misclassification rate drops sharply as the dimensions increase but stabilizes at 20 PCs (for GPA method) or harmonics two to seven (for the EFA method). For example, the first 20 PCs explain 98·8% of the variance, whereas harmonics two to seven explain 95·2% of the variance. Misclassification, however, increases after reaching *c*. 60 PCs [Fig. 3(a)] or after 15 harmonics [Fig. 3(b)].

COMPARISON OF CLASSIFICATION MODELS BASED ON GEOMETRIC MORPHOMETRICS AND SHAPE INDICES

Performance of LDA models built from geometric morphometrics and shape indices were compared [Fig. 4(a)]. Geometric morphometric methods give an overall mean accuracy of 94·5% (GPA) and 95·9% (EFA), compared with the shape indices method (90%), after optimization for dimension. A detailed breakdown on the performance measures by species is shown in Fig. 4(b). The performance of the different models, however, varies according to species. For example, the shape indices method performs slightly better in the identification of *P. maculatus*, *O. ruber* and *P. microdon*

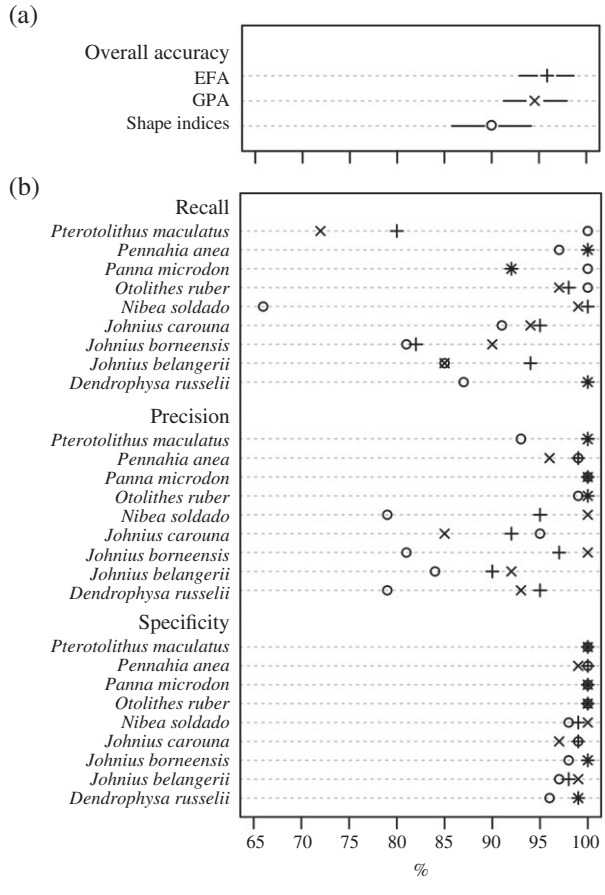

FIG. 4. Dot plot of classification performance (%) of models derived from the three featured extraction methods. (a) Comparison of overall accuracy (mean ± S.D.) for (+) elliptical Fourier analysis (EFA), (×) generalized Procrustes analysis (GPA) and (○) shape indices and (b) detailed breakdown of recall, precision and specificity for each of nine sciaenid species.

[Fig. 4(b)]. Overall, the geometric morphometric methods still perform better than the shape indices method and have a more balanced performance across all species. The aggregated LDA model gives an almost identical performance for the compared methods.

COMPARISON OF CLASSIFICATION PERFORMANCE FROM DIFFERENT IMAGE VIEWS

The classification performance using three different image views and their combinations are compared (Fig. 5). The proximal view performs best when evaluated alone, while the anterior and ventral views are less efficient. Overall accuracy increases when combinations of different views were used for training using the shape indices method, as compared with the GPA and EFA methods. When combinations of image views are used, the performance of the shape indices method is improved and is on par with

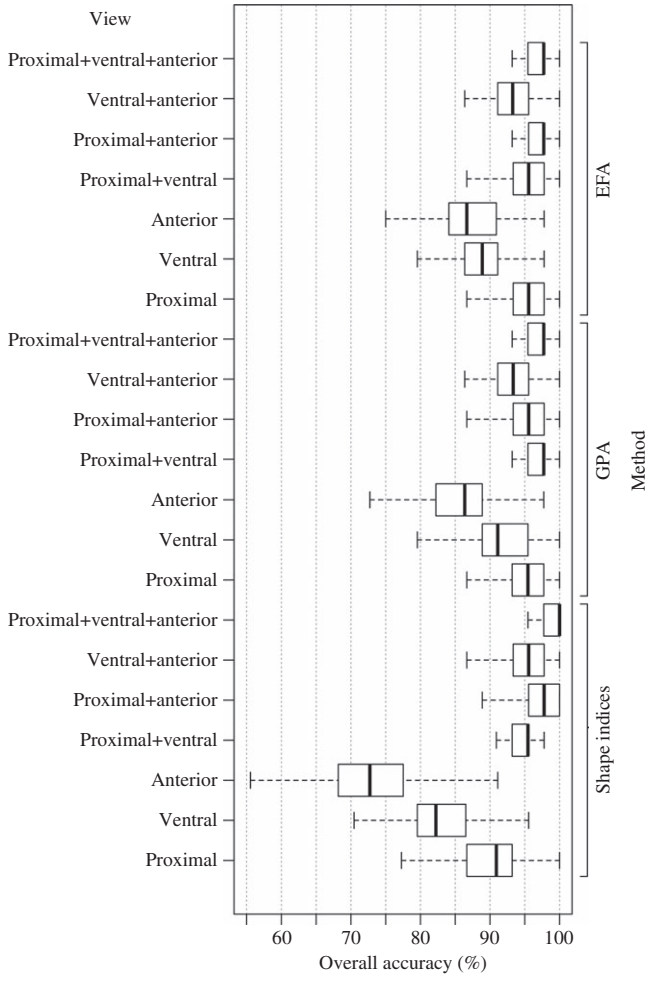

FIG. 5. Box-whisker plots of classification performance of elliptical Fourier analysis (EPA), generalized Procrustes analysis (GPA) and shape indices models derived from images of different aspects (proximal, ventral and anterior) of otoliths, and their combinations, expressed as distribution of overall accuracy (%). Middle bold line of box indicates median, box end indicates first and third quartiles and whiskers indicate range. Outliers are not included in the plot.

that of the geometric morphometric method, or even better in some cases. Among the pair-wise combinations, proximal-ventral and proximal-anterior combinations perform similarly well, but the ventral-anterior combination consistently performs worst for all the three methods.

EFFECT OF REJECTING UNCERTAIN PREDICTIONS

In the non-aggregated LDA model, the thresholding effect on posterior probability takes place only after a threshold value $\geq 0{\cdot}5$, where the accepted accuracy starts to increase while the accepted number of predicted specimens decreases [Fig. 6(a), (c), (e)]. The rate of drop in prediction number is faster than the increase in accuracy. This

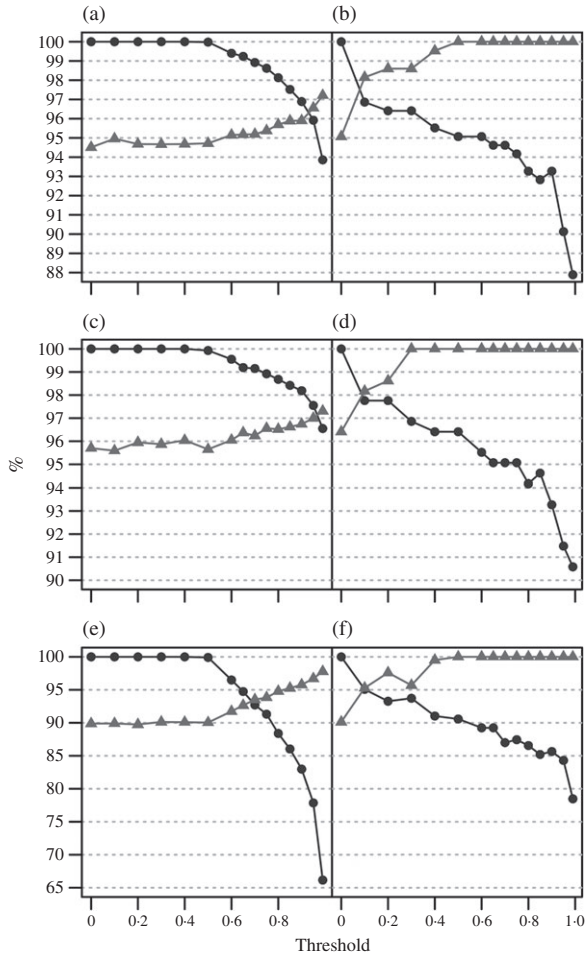

FIG. 6. Thresholding effects of posterior probability and proportion of votes on the overall accuracy and total reported predictions for (a, c, e) non-aggregated models and (b, d, f) aggregated models. Both models are linear discriminant analysis (LDA) models based on the (a, b) generalized Procrustes analysis (GPA), (c, d) elliptical Fourier analysis (EFA) and (e, f) shape indices methods. Overall accuracy and total prediction for (a, c, e) are means over all surrogate models from iterated *k*-folds cross-validations. All values of overall accuracy were calculated by excluding the uncertain predictions. A zero (0) threshold value means no thresholding and serves as the before threshold scenario for comparison. —●—, accepted prediction; —▲—, accepted accuracy.

suggests that some correct predictions were also excluded in the process. This pattern of response to thresholding is similar across all shape description methods, although it is more pronounced in the shape indices method.

In contrast, the aggregated model shows a much more sensitive response and its accepted accuracy increases to 100% at threshold values >0·5 [Fig. 6(b), (d), (f)]. More importantly, the drop in number of total accepted prediction is accompanied by a nearly equal increase in accepted accuracy. For example, the mean accepted accuracy for shape indices method [Fig. 6(f)] at a threshold of 0·5 increases by 9% from 91·0 to 100%, while the accepted prediction decreases by 9·7% from 100 to 90·3%. This

TABLE I. (a) Multivariate regression of sagittal otolith shape against otolith size, performed separately for each of the nine sciaenid species. The shape information (Procrustes distances) and centroid size were obtained from GPA performed on sliding semi-landmarks. (b) MANCOVA of effect of species on otolith shape differences using logarithm of centroid size of otolith as covariate

	d.f.*	*F*	*P*	r^2
(a) Within species				
Dendrophysa russelii	1, 29	2·64	<0·05	0·09
Johnius belangerii	1, 28	3·85	<0·05	0·12
Johnius borneensis	1, 14	2·56	<0·05	0·16
Johnius carouna	1, 31	0·94	>0·05	0·03
Nibea soldado	1, 20	10·2	<0·001	0·35
Otolithes ruber	1, 40	13·74	<0·001	0·26
Panna microdon	1, 12	0·34	>0·05	0·03
Pennahia anea	1, 35	44·84	<0·001	0·57
Pterotolithus maculatus	1, 5	0·4	>0·05	0·09
(b) Among species				
Species	8	71·89	<0·001	0·71
Log (centroid size)	1	18·67	<0·01	0·02
Total	222			

*For within-species tests, d.f. refers to that of regression and total shape variability respectively.

is an example of an ideal scenario for thresholding with a nearly equal trade-off. In contrast, for the non-aggregated model [Fig. 6(e)], the mean accepted accuracy increases by 5·3% from 89·8 to 95·2% at a threshold of 0·85, but at the expense of rejecting 14·2% of specimens (*i.e.* total accepted prediction decreases from 100 to 85·8%).

ALLOMETRIC EFFECT

Of all the species tested, otolith size (centroid size) of *P. microdon*, *N. soldado* and *O. ruber* show a highly significant effect on otolith shape ($P < 0{\cdot}001$), with considerable amount of within-species shape variations ($r^2 = 0{\cdot}26–0{\cdot}57$) explained by size [Table I(a)]. *Johnius belangerii*, *D. russelii* and *J. borneensis* also show significant otolith size effect ($P < 0{\cdot}05$). With all specimens combined, species differences account for most of the otolith shape variation (71%), while otolith size accounts for just 2% of the variation [Table I(b)].

Visualization of otolith shape differences due to otolith size is provided on a PCA plot based on three species that give highly significant results (Fig. 7). Small otoliths belonging to *O. ruber*, *N. soldado* and *P. microdon* are found to be similar [Fig. 7(a)], but the shapes of the larger otoliths are more variable between species. In all three species, the smaller otolith has a more rounded shape than the larger otolith, which is more elongate [Fig. 7(b)].

EVALUATION OF SEARCH AND PREDICT FUNCTIONS

The simulation result shows that the search algorithm for determining aspect and direction of the otoliths is correct in 98·7% of the cases, if the training images are used.

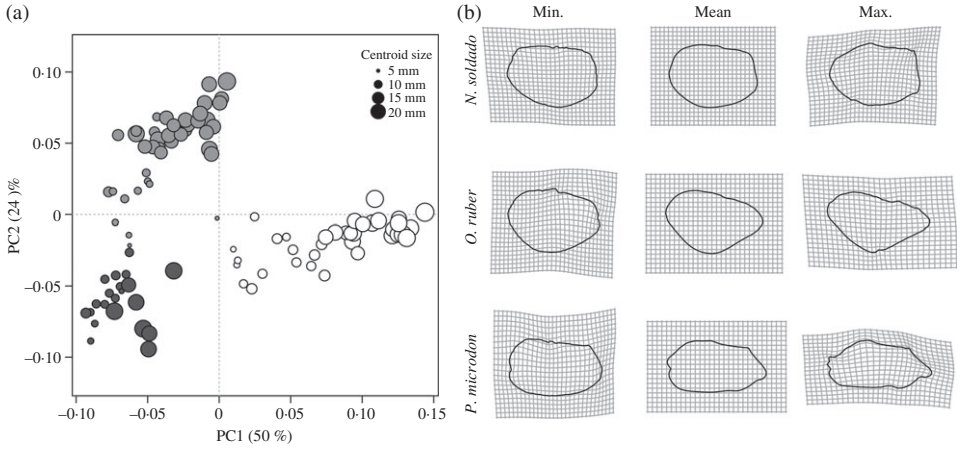

FIG. 7. (a) PCA bubble plot of otoliths of the top three sciaenid species showing shape changes due to allometry (obtained from generalized Procrustes analysis, GPA). Bubble size is relative to otolith size (centroid size; ■, *Nibea soldado*; ■, *Otolithes ruber*; □, *Panna microdon*). (b) Shape outlines of the smallest (min), mean shape (mean) and largest (max) otolith of *N. soldado*, *O. ruber* and *P. microdon*. Shape of the smallest and largest otolith is displayed as deformation from the mean shape, using the thin-plate spline method. PC, principal component.

Evaluation of the test images set show three mistakes (or 88% correct) at guessing the direction or side of the otolith. An example of an output of correct search results is illustrated in Fig. 8. The overall accuracy of classification performance for independent test images (Table II) is slightly lower than the results obtained from cross-validation. Results from the test set, however, give similar conclusions for relative performance. For the single aspect, the EFA method performs the best among the three methods. For two combined aspects, the EFA method also performs the best (especially the aggregated LDA) using combined proximal and ventral aspects. In all three methods, however, combining all three aspects only provides a marginal increase in classification performance. Better performance is achieved using the fixed test set as compared with the random test set (only proximal aspect tested) for both GPA and EFA methods, the probable reason being due to the incorrect guesses made by the search algorithm. On the other hand, the shape indices method seems to perform better in the random test set. Generally, the aggregated model performs slightly better than the normal LDA model.

DISCUSSION

DIMENSION REDUCTION

Dimension reduction in geometric morphometric methods not only helps to build a more simplified model by using fewer variables (*e.g.* from 200 variables of Cartesian co-ordinates to *c.* 20–30 variables in this study), but also helps to increase the performance of the classification model. The results show that a sufficient description of the otolith outline could be achieved despite the reduced number of dimensions,

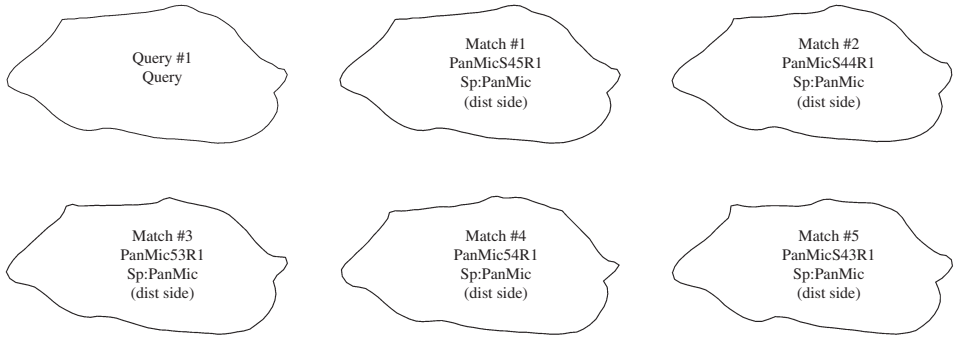

FIG. 8. Example of search result, displaying the outline of the query and its matches, ranked by Procrustes distance. Query #1 is the left otolith showing the proximal aspect, matched against configuration Fig. 2(d). dist side, distal side.

which also enhance classification performance. This can be attributed to the removal of noise in shape variations that is not important to classification in the higher order PCs or harmonics that contain shape information of minute changes (Claude, 2008). The present system has provided an objective way to evaluate dimension reduction for both GPA and EFA methods, with the main purpose of improving classification performance rather than achieving the best outline description (Baylac & Frieß, 2005). The present results show that both too little and too many shape dimensions could decrease the classification performance, consistent with the findings of previous studies (Reig-Bolaño *et al.*, 2011 for harmonics reduction; Chiari & Claude, 2012 for PC reduction). Dimension reduction can also be done at the stage of outline sampling by reducing the number of sampled semi-landmarks. The number of points sampled, however, has much less effect on EFA classification performance than the harmonics number (Reig-Bolaño *et al.*, 2011). The importance of selecting the right level of details for shape description has also been emphasized for other methods such as the wavelet transformation method (Sadighzadeh *et al.*, 2014). It is recommended that dimension reduction should be carried out to optimize the number of PCs and harmonics in GPA and EFA methods, respectively, especially for geometric morphometric data where the number of shape variables is likely to be more than the number of specimens (Chiari & Claude, 2012).

SHAPE DESCRIPTION METHOD

The present study shows that EFA is the shape description method with the best performance based on cross-validation evaluation and the test-image set. Therefore, EFA is recommended as the method of choice out of the three methods evaluated in the present study. The shape indices method has certain advantages over geometric morphometric methods, such as usage of dimensionless quantities that are inherently scale-invariant and rotation-invariant. No further normalization is required for the shape indices method and direct prediction can be made even with otoliths facing various directions. Although only 12 shape indices were used in this study, other shape indices can be added to increase the efficiency in shape description, such as those proposed by Tuset *et al.* (2006).

TABLE II. Classification performance based on independent test images ($n = 25$) for nine sciaenid species' sagittal otoliths

		Type of test images and classification model			
		Fixed side and direction		Random side and direction	
Description method	Otolith aspect	LDA model	Aggregated LDA model	LDA model	Aggregated LDA model
EFA	Proximal	92	92	88	88
GPA	Proximal	84	84	80	84
Shape indices	Proximal	76	84	84	88
EFA	Anterior	68	68	N/A	N/A
GPA	Anterior	52	52	N/A	N/A
Shape indices	Anterior	68	76	N/A	N/A
EFA	Ventral	80	84	N/A	N/A
GPA	Ventral	84	84	N/A	N/A
Shape indices	Ventral	72	76	N/A	N/A
EFA	Proximal + anterior	88	88	N/A	N/A
GPA	Proximal + anterior	84	80	N/A	N/A
Shape indices	Proximal + anterior	88	88	N/A	N/A
EFA	Proximal + ventral	92	96	N/A	N/A
GPA	Proximal + ventral	88	88	N/A	N/A
Shape indices	Proximal + ventral	88	88	N/A	N/A
EFA	Anterior + ventral	84	84	N/A	N/A
GPA	Anterior + ventral	88	88	N/A	N/A
Shape indices	Anterior + ventral	92	92	N/A	N/A
EFA	All 3 aspects	88	88	N/A	N/A
GPA	All 3 aspects	92	92	N/A	N/A
Shape indices	All 3 aspects	88	92	N/A	N/A

Results shown are overall accuracy (%). EFA, elliptical Fourier analysis; GPA, generalized Procrustes analysis; LDA, linear discriminant analysis; N/A, not evaluated.

The present study used three shape description methods, but other methods are available, *e.g.* wavelet transformation and curvature scale space representation (Parisi-Baradad *et al.*, 2005) and the different variations of Fourier transform (Reig-Bolaño *et al.*, 2010). Furthermore, shape description methods could be combined to increase classification performance, such as the combined geometric morphometric and linear measurement methods (Stransky *et al.*, 2008; Ginter *et al.*, 2012), or the combined geometric morphometric method with shape indices (Galley *et al.*, 2006; Viscosi *et al.*, 2010).

COMBINATION OF DIFFERENT IMAGE VIEWS

Literature shows that the proximal view has been used extensively (and almost exclusively) for shape analysis of the otolith (Monteiro *et al.*, 2005; Ponton, 2006; Tuset *et al.*, 2006; Sadighzadeh *et al.*, 2014). The present study shows that the proximal view indeed provides the best classification performance when only a single otolith aspect is considered. If further improvement of classification performance is desired,

the strategy of combining different image views of the otolith should be considered. The extra effort, however, of taking more images of the different aspects should be taken into consideration. The proximal-ventral view is recommended if a combination of two otolith aspects is desired. Chen *et al.* (2008) showed that more shape information could be derived by combining multiple 2D images to improve species identification. The present study has shown that by using a simple method to integrate the data from multiple views, classification performance could be enhanced. Combination of multiple views is particularly recommended for the shape indices method.

MODEL AGGREGATION AND PREDICTIVE CONFIDENCE

The use of the model aggregation technique has two main purposes: to train a more stabilized model that could potentially be better generalized and to improve the performance of rejecting uncertain predictions. Model aggregation has been suggested for classification if the training set is of small sample size because models trained from a small data set generally give higher variances (Beleites & Salzer, 2008). Conversely, a non-aggregated model is suggested for large data sets to increase computational speed (Ye *et al.*, 2011). In rejecting uncertain predictions, this study shows that the aggregated model performs better than the non-aggregated model, in agreement with the results of Beleites & Salzer (2008). It becomes almost inevitable, however, that some true positives are rejected in the process. For example, in the study by Ye *et al.* (2011), the total accepted prediction of 67% at a confidence level of 0·80 represents a drop from 98% at a confidence level of 0·10. The thresholding process, however, increased the accepted accuracy from 69 to 89%. The trade-off is a 20% increase of accepted accuracy at a cost of 32% of prediction numbers. Nevertheless, the approach of rejecting uncertain predictions by thresholding is still a useful strategy to single out potentially wrong predictions for further analysis, such as manual re-identification by the taxonomist. This process represents the semi-automatic approach that has been recommended to overcome limitations in automated identification systems (Grosjean *et al.*, 2004; Ye *et al.*, 2011). In the otolith case, manual re-identification could be performed with improved classification or by a closer examination of otolith features, such as the sulcus acusticus (Tuset *et al.*, 2008).

The present system uses posterior probability and surrogate model votes for thresholding, but there are other thresholding methods, such as those proposed by Ye *et al.* (2011). The LDA model is the classification model used in the present study and is the most common classifier used in the literature for biological classification. Many other types of classifiers, however, could be considered as alternatives (Mercier *et al.*, 2011; Wong *et al.*, 2014), including the use of combined classifiers (Hothorn & Lausen, 2005).

ALLOMETRIC EFFECT

The allometric effect is a problem of concern in otolith classification (Tuset *et al.*, 2006) because it contributes to a large amount of within-species variation. The shape variations could be so large that it has been used to discriminate fish age for some species (Beyer & Szedlmayer, 2010). Allometric effect on otolith shape due to size is significant for sciaenids of Brazil, although the explained variation was <25% (Monteiro *et al.*, 2005). Most of the sciaenid species examined in the present study also show

a similar range of explained variation, except *N. soldado* (35%) and *P. anea* (57%), which show very strong allometric effect due to size. Allometry, however, does not pose a serious challenge to classification performance in the present study as species still account for most of the shape variations, as shown by both the PCA bubble plot and MANCOVA result. Shape of the smaller otoliths appears more similar across species compared with the larger otoliths, while the latter tend to have more accentuated protrusions, similar to Mediterranean anguillids (Capoccioni *et al.*, 2011). The allometric effect should be tested and removed where possible, before classifying the otoliths. For example, Claude's (2013) allometric adjustment implementation can be easily adopted into the workflow. Nonetheless, allometric adjustment may not necessarily give a better classification performance, in which case, form-space rather than shape-space has been suggested for shape description (Ibáñez & O'Higgins, 2011; Claude, 2013).

SEARCH AND PREDICT FUNCTIONS

The present system consists of two components dealing with new image query. The search function works as a CBIR system that finds the most similar otolith configuration in the database by performing a dissimilarity ranking. Procrustes distance has been used as a measure of shape variation between specimens in the shape analysis (Adams *et al.*, 2013) and is proposed in the present study as a suitable metric to rank shape dissimilarity. For semi-landmark based configurations, a sliding procedure should be performed. Due to the concern of computation speed, however, the search function in the present study computes the Procrustes distance using non-sliding landmarks. The use of mean shape of each species helps to restrict the search range and shortens the computational time, but it decreases the matching accuracy. This trade-off becomes a greater problem if there is a high heterogeneity of among-species variance of the Procrustes distance to each species' mean shape. Nonetheless, the algorithm to restrict the search range is important. The AFORO system used an iterative algorithm to gradually restrict the search number from coarse to fine scale (Parisi-Baradad *et al.*, 2010). Both AFORO and the present system allow the visual comparison of outlines of query and matches so that the end user can determine whether there are any similar otoliths in the database. This also complements the predict function, which provides no indication of whether the query image is totally different from the species in the training database. One important feature of the present system, however, is that no prior knowledge of right or left otolith, nor direction of the queried image, is required by the end user to search for a match. This is an advantage to the inexperienced researcher, but this feature is not yet available in the current system for the combined aspects.

AUTOMATED CLASSIFICATION SYSTEM

Few taxonomic studies use the automated identification system for classification, thus its wider application is currently limited. Indeed, the potential applications of automated classification systems have not been fully explored and more developments are expected to aid fishery research (Fischer, 2013). A successful automated classification system should not only provide ease of use to end users (especially the non-specialist), but also flexibilities to optimize the system for advanced users (*e.g.* fishery scientists).

A good example of such kinds of flexibility is the ZooImage system developed for zooplankton classification (Grosjean *et al.*, 2004; Bell & Hopcroft, 2008).

In the present study, a generic automated classification system based on otoliths is developed. The system is built on the open-source platform of R, which provides constantly growing resources for morphometric analyses (Claude, 2008, 2013; Adams & Otárola-Castillo, 2013; Dryden, 2013; Bonhomme *et al.*, 2014; Schlager, 2014). Although this generic system is still under development, its future direction includes the development of a graphic user interface.

FURTHER APPLICATIONS

Otoliths have been used for various applications in fishery research. For discrimination of species of fishes, otolith shape may be useful to differentiate between morphologically similar species. For example, *J. carouna* and *J. belangerii* have similar morphological descriptions except for their body colouration (Sasaki, 2001), but their otoliths can be differentiated with high recall rate by the present identification system. As for the discrimination of fish stock populations, detailed shape variations may be required and the combination of multiple views will be more useful. Although the combined usage of all three otoliths (sagitta, lagena and utricula) in an individual has been shown to improve classification performance compared to using only the sagitta (Schulz-Mirbach & Plath, 2012), not all three otoliths can be easily recovered or identified for certain applications. For example, trophodynamic studies of marine mammals and birds that examine prey fishes in stomach contents or faeces (Bowen, 2000; Radhakrishnan *et al.*, 2010) may not find all three otoliths, or the non-specialist may not readily tell which side or direction of otolith is scanned. A quick search using the images of ingested or egested sagittae will be useful without such prior knowledge.

In conclusion, among the shape description methods, EFA gives the best classification performance, while the shape indices method remains a good alternative since it is simple and quick. A combination of multiple-shape description features, however, may further improve classification performance. At the image acquisition stage, if more detailed shape descriptions are needed, images of multiple aspects of the otolith are recommended. Dimension reduction helps in improving the classification performance of the geometric morphometric methods and the focus should be on reducing PCs–harmonic numbers. If the classification model is found to be unstable, the ensemble method such as model aggregation should be considered. Model aggregation is recommended if rejecting uncertain predictions is desired. The current version of the classification system allows researchers to build their own database and optimize the identification system that can be customized to other organisms and analyses. Although this classification system cannot provide the functionality of a large-scale centralized database, the aim is to further develop the system that will allow exchanges and merging of data sets to run in a local environment.

We thank the University of Malaya for providing research facilities and research grants RP008-2012C (V.C.C.), RG191-12SUS (K.H.L.). We appreciate T. F. Khang (University of Malaya) for initial discussion on the available methods and S. Schlager (University of Freiburg) for modifying his package to suit our 2D GM analysis. We thank S. L. Lee for providing the fish samples and S. M. Yeoh for processing fish samples and extracting otoliths. We are grateful to J. Claude (Université Montpellier 2) and one anonymous reviewer for their helpful comments on the manuscript.

Supporting Information

Supporting Information may be found in the online version of this paper:

Fig. S1. The proximal, anterior and ventral view of the sagitta of Sciaenidae. (a) *Dendrophysa russelii*; (b) *Johnius belangerii*; (c) *Johnius carouna*; (d) *Johnius borneensis*; (e) *Nibea soldado*; (f) *Otolithes ruber*; (g) *Pennahia anea*; (h) *Panna microdon*; (i) *Pterotolithus maculatus*; A, anterior; P, posterior; D, dorsal; V, ventral; Pr, proximal; Dl, distal.

Fig. S2. Example of semi-landmark sampling using the comb method. The major principal axis of the outline co-ordinates is first defined (horizontal line connecting the two red points). The axis is then divided into 50 equally spaced intervals, and perpendicular lines are drawn. Together, they form the comb (vertical dotted lines). The rest of the semi-landmarks are sampled where the comb and otolith outline intersect (vertical dotted lines). This sampling scheme, however, gives erroneous results on the folded outline (*).

Fig. S3. Schematic flowchart diagram of the conceptual framework used in the present study. (a) The work flow of building up a database of otolith shape information, using different shape description methods. (b) The strategies used to search and predict new specimens. The scheme shown is for the generalized Procrustes analysis (GPA) and elliptical Fourier analysis (EFA) methods since in the shape indices method, even if side and direction of otolith are unknown, prediction can still be made without performing a search first. LDA, linear discriminant analysis.

Table SI. List of sciaenid species otoliths studied, range of fish total length (L_T) and mass (M), and range of otolith length (L_O), height (H_O) and mass (O_M)

References

Adams, D. C. & Otárola-Castillo, E. (2013). Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution* **4,** 393–399.

Adams, D. C., Rohlf, F. J. & Slice, D. E. (2013). A field comes of age: geometric morphometrics in the 21st century. *Hystrix, the Italian Journal of Mammalogy* **24,** 7–14.

Baylac, M. & Frieß, M. (2005). Fourier descriptors, Procrustes superimposition and data dimensionality: an example of cranial shape analysis in modern human populations. In *Modern Morphometrics in Physical Anthropology* (Slice, D. E., ed), pp. 145–165. New York, NY: Kluwer Academic–Plenum Publishers.

Beleites, C. & Salzer, R. (2008). Assessing and improving the stability of chemometric models in small sample size situations. *Analytical and Bioanalytical Chemistry* **390,** 1261–1271.

Bell, J. L. & Hopcroft, R. R. (2008). Assessment of ZooImage as a tool for the classification of zooplankton. *Journal of Plankton Research* **30,** 1351–1367.

Beyer, S. G. & Szedlmayer, S. T. (2010). The use of otolith shape analysis for ageing juvenile red snapper, *Lutjanus campechanus*. *Environmental Biology of Fishes* **89,** 333–340.

Bonhomme, V., Picq, S., Gaucherel, C. & Claude, J. (2014). Momocs: outline analysis using R. *Journal of Statistical Software* **56,** 1–24.

Bookstein, F. L. (1997). Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis* **1,** 225–243.

Bouckaert, R. R. (2003). Choosing between two learning algorithms based on calibrated tests. In *Proceedings of the Twentieth International Conference (ICML 2003) on Machine Learning* (Fawcett, T. & Mishra, N., eds), pp. 51–58. Washington, DC: AAAI Press.

Bowen, W. D. (2000). Reconstruction of pinniped diets: accounting for complete digestion of otoliths and cephalopod beaks. *Canadian Journal of Fisheries and Aquatic Sciences* **57,** 898–905.

Breiman, L. (1996). Bagging predictors. *Machine Learning* **24,** 123–140.

Capoccioni, F., Costa, C., Aguzzi, J., Menesatti, P., Lombarte, A. & Ciccotti, E. (2011). Ontogenetic and environmental effects on otolith shape variability in three Mediterranean European eel (*Anguilla anguilla*, L.) local stocks. *Journal of Experimental Marine Biology and Ecology* **397,** 1–7.

Cavalcanti, M. J., Monteiro, L. R. & Lopes, P. R. (1999). Landmark-based morphometric analysis in selected species of serranid fishes (Perciformes: Teleostei). *Zoological Studies* **38,** 287–294.

Chen, H., Bart, H. L. Jr. & Huang, S. (2008). Integrated feature selection and clustering for taxonomic problems within fish species complexes. *Journal of Multimedia* **3,** 10–17.

Chiari, Y. & Claude, J. (2012). Morphometric identification of individuals when there are more shape variables than reference specimens: a case study in Galápagos tortoises. *Comptes Rendus Biologies* **335,** 62–68.

Chong, V. C., Lee, P. K. Y. & Lau, C. M. (2010). Diversity, extinction risk and conservation of Malaysian fishes. *Journal of Fish Biology* **76,** 2009–2066. doi: 10.1111/j.1095-8649.2010.02685.x

Claude, J. (2008). *Morphometrics with R*. New York, NY: Springer.

Claude, J. (2013). Log-shape ratios, Procrustes superimposition, elliptic Fourier analysis: three worked examples in R. *Hystrix, the Italian Journal of Mammalogy* **24,** 94–102.

Corti, M. & Crosetti, D. (1996). Geographic variation in the grey mullet: a geometric morphometric analysis using partial warp scores. *Journal of Fish Biology* **48,** 255–269.

Eschmeyer, W. N., Fricke, R., Fong, J. D. & Polack, D. A. (2010). Marine fish diversity: history of knowledge and discovery (Pisces). *Zootaxa* **2525,** 19–50.

Fadda, C., Faggiani, F. & Corti, M. (1997). A portable device for the three dimensional landmark collection of skeletal elements of small mammals. *Mammalia* **61,** 622–627.

Fischer, J. (2013). Fish identification tools for biodiversity and fisheries assessments: review and guidance for decision-makers. *FAO Fisheries and Aquaculture Technical Paper* **585**. Available at http://www.fao.org/docrep/019/i3354e/i3354e00.htm/

Galley, E. A., Wright, P. J. & Gibb, F. M. (2006). Combined methods of otolith shape analysis improve identification of spawning areas of Atlantic cod. *ICES Journal of Marine Science* **63,** 1710–1717.

Ginter, C. C., DeWitt, T. J., Fish, F. E. & Marshall, C. D. (2012). Fused traditional and geometric morphometrics demonstrate pinniped whisker diversity. *PLoS One* **7,** e34481. doi: 10.1371/journal.pone.0034481

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society B* **53,** 285–339.

Grosjean, P., Picheral, M., Warembourg, C. & Gorsky, G. (2004). Enumeration, measurement and identification of net zooplankton samples using the ZOOSCAN digital imaging system. *ICES Journal of Marine Science* **61,** 518–525.

Guisande, C., Manjarrés-Hernández, A., Pelayo-Villamil, P., Granado-Lorencio, C., Riveiro, I., Acuña, A., Prieto-Piraquive, E., Janeiro, E., Matías, J. M., Patti, B., Mazzola, S., Jiménez, S., Duque, V. & Salmerón, F. (2010). *IPez*: an expert system for the taxonomic identification of fishes based on machine learning techniques. *Fisheries Research* **102,** 240–247.

Herler, J., Kerschbaumer, M., Mitteroecker, P., Postl, L. & Sturmbauer, C. (2010). Sexual dimorphism and population divergence in the Lake Tanganyika cichlid fish genus *Tropheus*. *Frontiers in Zoology* **7,** 4.

Hothorn, T. & Lausen, B. (2005). Bundling classifiers by bagging trees. *Computational Statistics & Data Analysis* **49,** 1068–1078.

Hu, J., Li, D., Duan, Q., Han, Y., Chen, G. & Si, X. (2012). Fish species classification by color, texture and multi-class support vector machine using computer vision. *Computers and Electronics in Agriculture* **88,** 133–140.

Ibáñez, A. L. & O'Higgins, P. (2011). Identifying fish scales: the influence of allometry on scale shape and classification. *Fisheries Research* **109,** 54–60.

Lombarte, A., Palmer, M., Matallanas, J., Gómez-Zurita, J. & Morales-Nin, B. (2010). Ecomorphological trends and phylogenetic inertia of otolith sagittae in Nototheniidae. *Environmental Biology of Fishes* **89,** 607–618.

Mercier, L., Darnaude, A. M., Bruguier, O., Vasconcelos, R. P., Cabral, H. N., Costa, M. J., Lara, M., Jones, D. L. & Mouillot, D. (2011). Selecting statistical models and variable combinations for optimal classification using otolith microchemistry. *Ecological Applications* **21,** 1352–1364.

Monteiro, L. R., Beneditto, A. P. M. D., Guillermo, L. H. & Rivera, L. A. (2005). Allometric changes and shape differentiation of sagitta otoliths in sciaenid fishes. *Fisheries Research* **74,** 288–299.

Parisi-Baradad, V., Lombarte, A., García-Ladona, E., Cabestany, J., Piera, J. & Chic, O. (2005). Otolith shape contour analysis using affine transformation invariant wavelet transforms and curvature scale space representation. *Marine and Freshwater Research* **56,** 795–804.

Parisi-Baradad, V., Manjabacas, A., Lombarte, A., Olivella, R., Chic, Ò., Piera, J. & García-Ladona, E. (2010). Automated taxon identification of teleost fishes using an otolith online database – AFORO. *Fisheries Research* **105,** 13–20.

Pau, G., Fuchs, F., Sklyar, O., Boutros, M. & Huber, W. (2010). EBImage – an R package for image processing with applications to cellular phenotypes. *Bioinformatics* **26,** 979–981.

Ponton, D. (2006). Is geometric morphometrics efficient for comparing otolith shape of different fish species? *Journal of Morphology* **267,** 750–757.

Radhakrishnan, K. V., Liu, M., He, W., Murphy, B. R. & Xie, S. (2010). Otolith retrieval from faeces and reconstruction of prey-fish size for great cormorant (*Phalacrocorax carbo*) wintering at the East Dongting Lake National Nature Reserve, China. *Environmental Biology of Fishes* **89,** 505–512.

Reig-Bolaño, R., Marti-Puig, P., Lombarte, A., Soria, J. A. & Parisi-Baradad, V. (2010). A new otolith image contour descriptor based on partial reflection. *Environmental Biology of Fishes* **89,** 579–590.

Reig-Bolaño, R., Marti-Puig, P., Gallego, E., Masferrer, G., Lombarte, A., Ferrer-Arnau, L. & Parisi-Baradad, V. (2011). Feature selection for analyzing and retrieving fish otoliths using Elliptic Fourier Descriptors of shapes. In *7th International Conference on Next Generation Web Services Practices (NWeSP), 19–21 October 2011*, pp. 290–295. Salamanca: IEEE. doi: 10.1109/NWeSP.2011.6088193

Rohlf, F. J. & Archie, J. W. (1984). A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Systematic Biology* **33,** 302–317.

Rohlf, F. J. & Slice, D. (1990). Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Biology* **39,** 40–59.

Sadighzadeh, Z., Valinassab, T., Vosugi, G., Motallebi, A. A., Fatemi, M. R., Lombarte, A. & Tuset, V. M. (2014). Use of otolith shape for stock identification of John's snapper, *Lutjanus johnii* (Pisces: Lutjanidae), from the Persian Gulf and the Oman Sea. *Fisheries Research* **155,** 59–63.

Sasaki, K. (1989). Phylogeny of the family Sciaenidae, with notes on its zoogeography (Teleostei: Perciformes). *Memoirs of the Faculty of Fisheries-Hokkaido University* **36,** 1–137.

Sasaki, K. (2001). Sciaenidae. In *FAO Species Identification Guide for Fishery Purposes. The Living Marine Resources of the Western Central Pacific*, Vol. 5, *Bony Fishes*, *Part 3* (*Menidae to Pomacentridae*) (Carpenter, K. E. & Niem, V. H., eds), pp. 3117–3174. Rome: FAO.

Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. (2012). NIH image to ImageJ: 25 years of image analysis. *Nature Methods* **9,** 671–675.

Schulz-Mirbach, T. & Plath, M. (2012). All good things come in threes – species delimitation through shape analysis of saccular, lagenar and utricular otoliths. *Marine and Freshwater Research* **63,** 934–940.

Sokolova, M. & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management* **45,** 427–437.

Stransky, C., Baumann, H., Fevolden, S.-E., Harbitz, A., Høie, H., Nedreaas, K. H., Salberg, A.-B. & Skarstein, T. H. (2008). Separation of Norwegian coastal cod and Northeast Arctic cod by outer otolith shape analysis. *Fisheries Research* **90,** 26–35.

Tuset, V. M., Rosin, P. L. & Lombarte, A. (2006). Sagittal otolith shape used in the identification of fishes of the genus *Serranus*. *Fisheries Research* **81,** 316–325.

Tuset, V. M., Lombarte, A. & Assis, C. A. (2008). Otolith atlas for the western Mediterranean, north and central eastern Atlantic. *Scientia Marina* **72,** 7–198.

Venables, W. N. & Ripley, B. D. (2002). *Modern Applied Statistics with S*, 4th edn. New York, NY: Springer.

Viscosi, V., Loy, A. & Fortini, P. (2010). Geometric morphometric analysis as a tool to explore covariation between shape and other quantitative leaf traits in European white oaks. In *Tools for Identifying Biodiversity: Progress and Problems* (Nimis, P. L. & Vignes, L. R., eds), pp. 257–261. Trieste: Edizioni Università di Trieste.

Wong, J. Y., Chen, H. N., Chan, B. K. K., Tan, I. K. P. & Chong, V. C. (2014). A combined morphological and molecular approach in identifying barnacle cyprids from the Matang Mangrove Forest Reserve in Malaysia: essentials for larval ecology studies. *Raffles Bulletin of Zoology* **62,** 317–329.

Ye, L., Chang, C.-Y. & Hsieh, C.-H. (2011). Bayesian model for semi-automated zooplankton classification with predictive confidence and rapid category aggregation. *Marine Ecology Progress Series* **441,** 185–196.

Electronic References

Dryden, I. L. (2013). *Shapes*: *Statistical shape analysis*. Available at http://CRAN.R-project.org/package=shapes/

Landini, G. (2008). Advanced shape analysis with ImageJ. In *Proceedings of the Second ImageJ User and Developer Conference*, 6–7 November 2008, pp. 116–121. Luxembourg: CRP Henri Tudor. Available at http://www.mecourse.com/landinig/software/software.html

Lin, C-H. & Chang, C-W. (2013). *Otolith Atlas of Taiwan Fishes*. Available at http://oto.biodiv.tw/

Schlager, S. (2014). *Morpho*: *Calculations and Visualisations Related to Geometric Morphometrics*. Available at: http://CRAN.R-project.org/package=Morpho/