Received: 11 January 2024 | Accepted: 26 June 2024

DOI: 10.1111/2041-210X.14388

Methods in Ecology and Evolution

APPLICATION

3DKMI: A MATLAB package to generate shape signatures from Krawtchouk moments and an application to species delimitation in planktonic foraminifera

Huahua Lin[1] | Wenshu Zhang[1,2] | James M. Mulqueeney[3,4] | Anieke Brombacher[4,5] | Alex Searle-Barnes[4] | Mark Nixon[1] | Xiaohao Cai[1] | Thomas H. G. Ezard[4]

[1]Electronics and Computer Science, University of Southampton, Southampton, UK

[2]Creative Computing Institute, University of the Arts London, London, UK

[3]Life Sciences, Natural History Museum, London, UK

[4]Ocean and Earth Science, University of Southampton Waterfront Campus, National Oceanography Centre Southampton, Southampton, UK

[5]Earth and Planetary Sciences, Yale University, New Haven, Connecticut, USA

Correspondence
Huahua Lin
Email: huahua.lin@soton.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: NE/P019269/1

Handling Editor: Gustavo Burin

Abstract

1. The rapid and repeatable characterization of individual morphology has advanced automated taxonomic classification. The most direct study of evolutionary processes is, however, not from taxonomic description, but rather of the evolution of the traits that comprise individuals and define species. Repeatable signatures of individual morphology are crucial for analysing the response to selection at scale, and thus tracking evolutionary trajectories through time and across species boundaries.
2. Here, we introduce our 3DKMI—an open-source MATLAB package designed for the study of morphology using three-dimensional (3D) Krawtchouk moment invariants. The volumetric features derived from the 3D images remain stable under translation, scaling and rotation and, for an image of size $128 \times 128 \times 128$ can be computed in less than 0.1 s.
3. We applied our package as a case study on a collection of 300 X-ray computed tomography scans of planktonic foraminifera specimens across five species to (1) assess the invariance of the features under different transformations and (2) analyse morphological differences among species based on the extracted characteristics.
4. We show that 3DKMI has the capacity to efficiently and repeatedly characterize the signatures of individual morphology. In the future, we hope that the 3D feature extraction technique 3DKMI will be widely applied to digital collections to advance research in ecology and evolution.

KEYWORDS
3D Krawtchouk moments, classification, computer vision, feature extraction, machine learning, planktonic foraminifera, shape, tomography

1 | INTRODUCTION

Feature extraction in computer vision has emerged as a powerful tool in the field of evolutionary biology, revolutionizing the way researchers study and analyse complex biological systems (Porto & Voje, 2020; Wäldchen & Mäder, 2018). By employing various algorithms and techniques, computer vision systems can provide a high-level understanding of visual data (Porto et al., 2021) to automatically detect and analyse morphological features, such as the shape (Cooney et al., 2017; Spampinato et al., 2010), size (Alsmadi et al., 2010; Hsiang et al., 2019), and colour patterns (Alsmadi et al., 2011; Cooney et al., 2019) of organisms. Through extracting and quantifying these features, researchers can conduct comparative analysis across different individuals, species, populations and communities, providing valuable insights into the evolutionary processes and adaptations that shape the diversity of life on Earth.

One common application of feature extraction is the problem of automatic classification of individuals to species (Fassnacht et al., 2016; Spampinato et al., 2010). The individuals grouped into self-similar clusters, and then often mapped back to manual taxonomic classifications as species, comprise distinctive sets of genetic (Pons et al., 2006; Talavera et al., 2013) or phenotypic traits (Pearson & Ezard, 2014). Bridging the gap from traditional taxonomy to automated computer vision pivots on the characterization of those sets of phenotypic traits that define species, but assessing which traits carry the most weight in determining species classifications is often subsumed within a "black box" rather than the core output of model classifications.

Deeper morphological understanding promises substantial potential to improve our understanding of phenotypic evolution (Felice et al., 2018). However, direct morphological evidence of how biological diversity came to exhibit its strikingly discontinuous pattern is not complete (Rolland et al., 2023). Rapid and repeatable methods for morphometric data extraction would therefore boost the study of phenotypic evolution. Geometric morphometric approaches have become the gold standard in evolutionary biology due to the limitations of linear methods in distinguishing complex morphologies (Viacava et al., 2023; Zelditch et al., 2012), providing fuller insights into how morphology relates to underlying functions (Fabre et al., 2014) and developmental processes (Goswami et al., 2014), while also offering less biased descriptors (Zelditch et al., 2012). Despite recent advances, such as enhancing shape characterization with sliding semi-landmarks (Bardua et al., 2019), manual work is susceptible to observer bias, lacks repeatability and is time-consuming. While using automated approaches including automated landmarking (Devine et al., 2020) and landmark-free methods (Mulqueeney et al., 2024) may help overcome these issues, they still require homology for either analysis or alignment, and thus become less effective as more disparate shapes are compared (Collyer et al., 2015).

Image moments, initially conceptualized by Hu (1962), serve as mathematical descriptors representing the spatial distribution of points in an image, widely employed to encapsulate global features such as an object's shape, thus eliminating the necessity for landmarking. Following this seminal work, a variety of new moments (Mukundan et al., 2001; Teague, 1980; Yap et al., 2003) based on different polynomials rapidly emerged. Among them, Krawtchouk moments stand out because the basis functions are orthogonal in the discrete domain of the image coordinate space, thus minimizing feature redundancy and eliminating the need for numerical approximation (Yap et al., 2003). Other advantages of using Krawtchouk moments for feature description over deep learning methods lie in their simplicity, efficiency and interpretability, providing a straightforward mathematical framework for representing image features without requiring comparisons of homologous points, extensive computational resources, or large amounts of labelled data.

Here, we describe our developed 3DKMI package, a MATLAB package designed for the automatic generation of shape features. Our package implements the code for computing three-dimensional (3D) Krawtchouk moment invariants, derived from discrete and orthogonal 3D Krawtchouk polynomials. These invariants can consistently describe the individual regardless of translation, scaling and rotation. Unlike previous studies (Benouini et al., 2018; Sit et al., 2019), our novel in-house experiment focuses on a 3D dataset of planktonic foraminifera species. To the best of our knowledge, we are the first to make this code publicly available (https://github.com/krawtchouk/Krawtchouk-Moment-Invariants) and apply it to a 3D ecological dataset, complemented by a quick-start guide and examples available at http://krawtchouk.github.io.

2 | PACKAGE FEATURES

3DKMI is written in MATLAB, version R2022b, with backward compatibility. The program's dependencies include the "Symbolic Math Toolbox" and the "Image Processing Toolbox". Below, we present the two primary functions—readImage and get3DKMI—within the package.

2.1 | readImage: Construct a 3D array

The function readImage takes two input parameters: the path where the 3D image data is stored along with the desired image dimensions dim in integer format. It produces a corresponding 3D array that represents the input image and is called with the following command:

```
readImage(path, dim)
```

The 3D image will be resized to the given desired dimensions dim so that it can be fitted into a cube formed by the desired dimensions without disrupting its geometry. It is important to use consistent desired dimensions dim when comparing the features extracted from different 3D images.

The path accommodates two distinct scenarios: (i) it can either point to a folder containing a sequence of two-dimensional (2D) image slices, where each is stored as for example an individual .png file or (ii) it can direct to a 3D image file. In the former scenario, the method processes these image slices sequentially, arranging them based on the

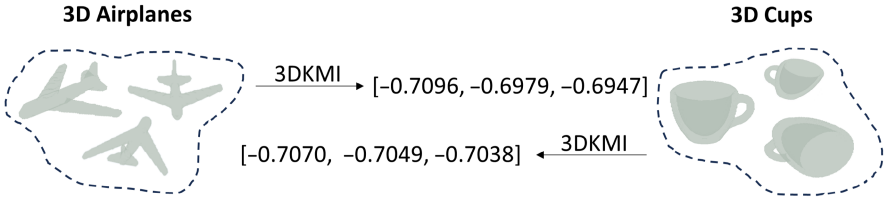

FIGURE 1 Two different types of three-dimensional (3D) shapes and their corresponding features represented by 3D Krawtchouk moments.

filename order (e.g. 0.png, 1.png, …, n.png). This aggregation results in an image stack, ultimately manifested as a 3D array representation. In the latter scenario, our current implementation supports reading .im and .stl formats for 3D image files and generating voxelized data. Reading .im files follows the process provided on McGill's 3D shape benchmark website (Siddiqi et al., 2008). For .stl files, the existing MATLAB built-in functions can only extract the vertex and face information from the image. Therefore, we employ a voxelization method to further derive the 3D array (Adam, 2023). If users find that their file types are unsupported, they can generate their own 3D arrays representing images and then directly use the get3DKMI function.

2.2 | get3DKMI: Feature extraction

The function get3DKMI takes a 3D array arr as input, along with the desired dimensions dim in integer format and produces 3D Krawtchouk moment invariants as output. The function is called using the following command:

get3DKMI(arr, dim, varargin)

where varargin allows the function to accept up to two input parameters: POI (point-of-interest) and/or the order of the 3D Krawtchouk moments. Notably, providing a POI as input implies that the user expects to extract localized features from the image, whereas the absence of this input results in the extraction of features from the image as a whole.

The distinction between local and global feature extraction lies in how weights are assigned to individual pixels within the image. The weight function in Yap et al. (2003) and Sit et al. (2019) was introduced to create a more stable set of polynomials, but we found it to be an integral part of extracting local features around the POI. In the case of local extraction, those pixels around the POI can hold greater significance than the rest of the region by using the weight function. The user can also select the moment's order. The order of moments denotes the mathematical degree of the moments being computed, with different orders capturing distinct aspects of an image's characteristics. A higher order yields a more detailed description of the local or global image, though it requires more computational resources. The length of the output invariants depends on the order of the 3D Krawtchouk moments.

The parameter dim is also included in get3DKMI in case the users themselves generate the 3D array that needs to be resized. Additionally, the spherical shape of the interested local region can be preserved when the image is in the 3D cubic formed by the desired dimensions dim. Once the image dimensions change, the relative positions of the POI will be recalculated accordingly to ensure the features are still extracted around the interested region. However, the size of the image influences the program's computation time, that is larger dimensions result in increased time spent on matrix computations. Our benchmark computation time is provided in Section 3.2. We suggest users experiment with various values according to their hardware capabilities. Detailed mathematical instructions for obtaining invariants through Krawtchouk moments can be found on our GitHub website http://krawtchouk.github.io.

2.3 | How 3DKMI works

To illustrate the functionality of 3DKMI, we selected two distinct 3D shapes from the McGill 3D shape benchmark (Siddiqi et al., 2008): an airplane and a cup as shown in Figure 1. We employed our package to extract their global features, and the calculated second-order invariant values are [−0.7096, −0.6979, −0.6947] for the airplanes and [−0.7070, −0.7049, −0.7038] for the cups. Despite the substantial dissimilarity between the shapes of the airplane and the cup, the invariant values exhibit a small absolute difference between the calculated invariant values of the two shapes, which is a characteristic of the 3D Krawtchouk moments. Note that, at the scale of the feature space generated by the 3D Krawtchouk moments, that small absolute difference is not small for 3DKMI and is sufficient to determine the difference between three airplane images and three cup images in Figure 1. The feature distance between the original airplane image and the transformed airplane image is smaller than that between the original airplane image and the cup image. Likewise, the values of the local features exhibit similar behaviour, allowing us to view the local region as a novel shape. In short, 3DKMI can represent each different 3D shape with a unique signature.

3 | RESULTS: APPLICATIONS ON SPECIES DIFFERENCES IN PLANKTONIC FORAMINIFERA

3.1 | Data

Modern high-resolution X-ray computed tomography (CT) makes it possible for us to observe and analyse the internal and external structures of planktonic foraminifera that are less than 1 mm in

diameter. We compiled a dataset of 300 individuals over five species from Ocean Drilling Program Site 925, located on the Ceara Rise in the Equatorial Atlantic, and Integrated Ocean Drilling Program Site U1313 within the North Atlantic gyre; for site-specific oceanographic settings, see Brombacher et al. (2021) and references therein. The five species are *Globigerinoidesella fistulosa*, *Menardella exilis*, *Truncorotalia crassaformis*, *Sphaeroidinella dehiscens* and *Trilobatus sacculifer*, with some sample images shown in Figure 2. Each species has a sample size of 50, except for *Menardella exilis*, which has an additional 50 samples for efficiency benchmarking. One of us (A.B.) generated taxonomic labels through key diagnostic features, including the shape and arrangement of the chambers, the shape and position of the apertures, and other morphological features of the shells. Each individual was randomly drawn, within the strata of its species classification, from a larger sample of 1226 specimens.

For each individual, approximately 300 2D image slices were obtained by CT scanning. As our emphasis was solely on morphology, we obtained binary volume representations of each specimen from these slices by applying some pre-processing steps such as cropping and segmentation (Figure 3). Global threshold values for segmentation are manually determined to remove background noise while preserving shape integrity. This step can be automated by using threshold values selected automatically by, for example Otsu's method (Otsu, 1975).

3.2 | Implementation efficiency

We utilized 100 specimens from the species *M. exilis* to benchmark the efficiency of 3DKMI in feature extraction, running on a laptop equipped with an AMD Ryzen 5 5600H CPU and 16 GB of RAM. For parameter selection, each image was resized to $128 \times 128 \times 128$ (i.e. dim is set to 128) and the order of moments was fixed to 2. The entire workflow after pre-processing for these 100 *M. exilis*, including image reading and feature extraction, took ~58 s, with the majority of the time dedicated to reading slices to form the 3D array data. The average runtime for each sample was 0.48 s for forming the 3D array and 0.09 s for calculating the 3D Krawtchouk moment invariants.

3.3 | Translation, scale and rotation invariance

In this experiment, we investigated the property of the 3D Krawtchouk moment invariants on the dataset under different geometric transformations, including translation, scaling, and rotation. The 3D Krawtchouk moment invariants were computed up to the second order. For comparison, the 3D images were resized to cubes of the same dimension, that is $128 \times 128 \times 128$. Both global and local features were tested. In our case, we selected the image's centroid as the POI.

One specimen from the species *M. exilis* was randomly selected for demonstration purposes. Three transformations were applied separately to analyse their impact on invariants. We first investigated the translation invariance by shifting the object within the cube with a step of (1, 1, 1) using translation vectors ranging from (−10, −10, −10) to (10, 10, 10). To assess the scale invariance, we applied scaling factors from 0.7 to 1.3 with a step size of 0.05 to the object. Finally, the rotation invariance was explored by rotating the object around each axis with rotation angles ranging from 0° to 90°, at intervals of 10°. To preserve the object's integrity, we scaled down any transformed object that exceeds the dimensions of the cube.

Figure 4 depicts the mean squared error (MSE) between the 3D Krawtchouk moment invariants of the original object and the ones under translation, scaling and rotation. The overall results in

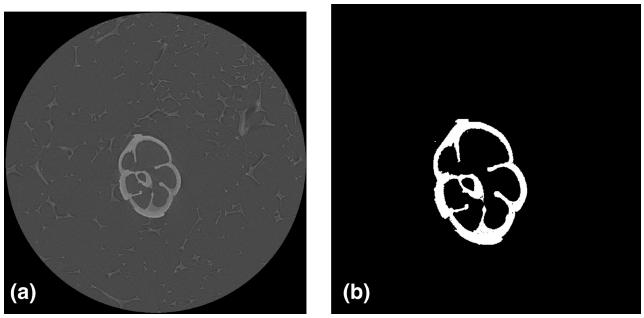

FIGURE 3 (a) A two-dimensional (2D) slice of a single *Menardella exilis* specimen reconstructed by X-ray computed tomography (Zeiss Reconstructor v14) and (b) the same 2D slice represented as a binary image after cropping and segmentation using a manually specified threshold value.

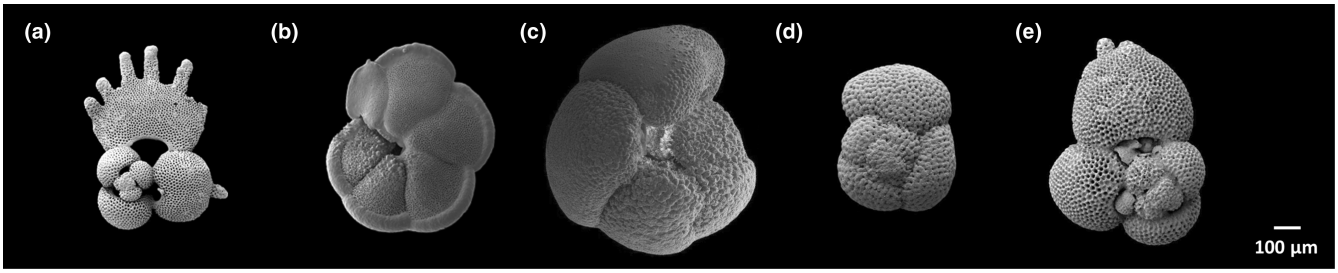

FIGURE 2 Example images of the five different species of planktonic foraminifera used in this application. (a) *Globigerinoidesella fistulosa* (Poole & Wade, 2019), (b) *Menardella exilis* (Woodhouse et al., 2023), (c) *Truncorotalia crassaformis* (Lam & Leckie, 2020a), (d) *Sphaeroidinella dehiscens* (Lam & Leckie, 2020b) and (e) *Trilobatus sacculifer* (Poole & Wade, 2019). The scalar bar for images is 100 µm.

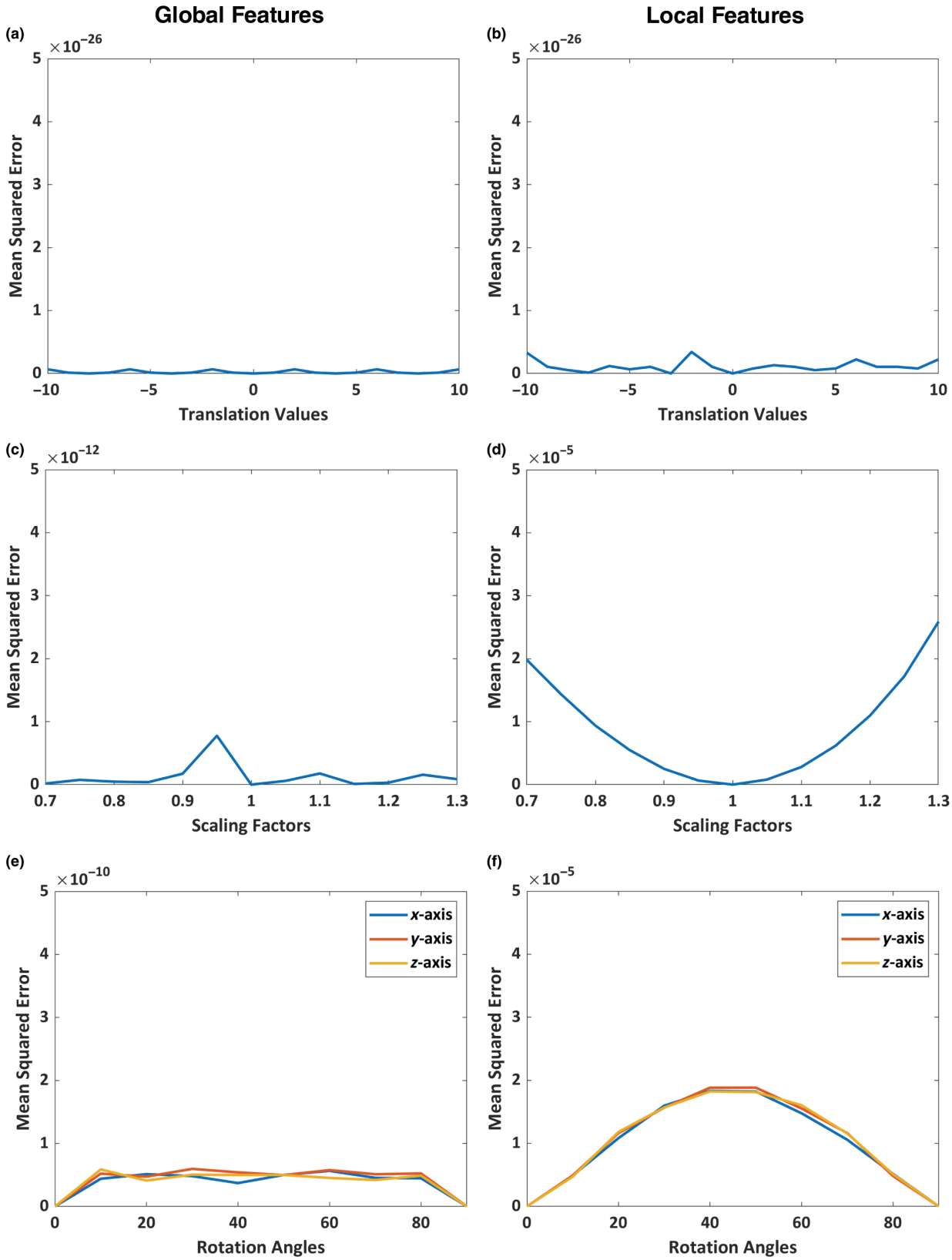

FIGURE 4 Invariance performance validation of 3DKMI. The left and right columns show the mean squared error between the original object and the transformed ones regarding global features and local features, respectively. Three transformations are involved in total: (a, b) when the object is shifted (translation); (c, d) when the size of the object changes (scale); and (e, f) when the object rotates about the *x*-axis, *y*-axis and *z*-axis (rotation). All transformed objects still fit the size of the original image. Note the different scales on the *y*-axes.

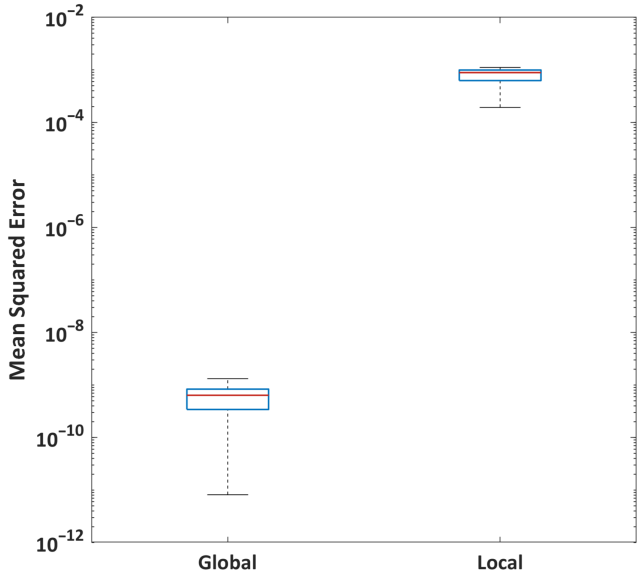

FIGURE 5 Invariance performance validation of 3DKMI when three random transformations (including translation, scaling, and rotation) are performed simultaneously. The mean squared error (MSE) between the 3D Krawtchouk moment invariants of the original object and the transformed one comes from 50 random transformations; the median MSE using global features is $6e^{-10}$ while the one using local features is $8e^{-4}$.

Figure 4 indicate consistently low errors regardless of the different geometric transformations applied to the given object. For the comparison between the two feature extraction manners, that is global and local, the error of the former is always smaller than that of the latter. In the global scaling and rotation experiments, the MSE values are relatively high compared to the translation experiments because when a 3D image is scaled and rotated, if the voxels in the transformed image are in non-integer positions, interpolation will be used to map the voxels in the original image to their new positions in the transformed image. This small change will slightly affect the extracted features, but the MSE values are still less than $1e^{-10}$. The experimental results also show that local feature extraction is most sensitive to scaling because when the object is scaled, the weight function remains to assign higher weights to the same size region of interest. Additionally, as shown in the right panel of Figure 4f, the error exhibits symmetry when the object is rotated from 0° to 90° around the axis, with the largest error occurring at 45° of rotation. This angle of rotation causes the object to exceed the cube dimension the most, and the object is thus at the same time scaled down the most to ensure its integrity.

Furthermore, we randomly and simultaneously applied the aforementioned three transformations to validate the invariance property of 3D Krawtchouk moment invariants under more

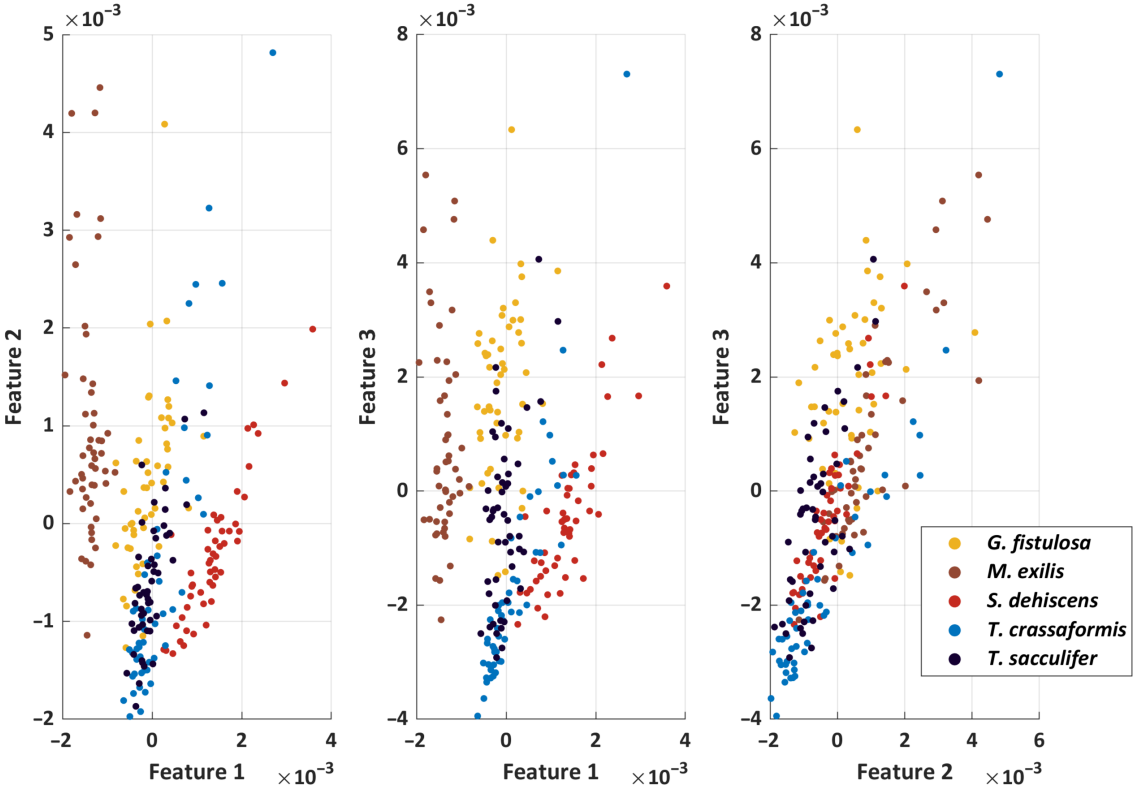

FIGURE 6 Two-dimensional projections on each axis of the Krawtchouk moment invariants for 250 specimens represented by 3DKMI.

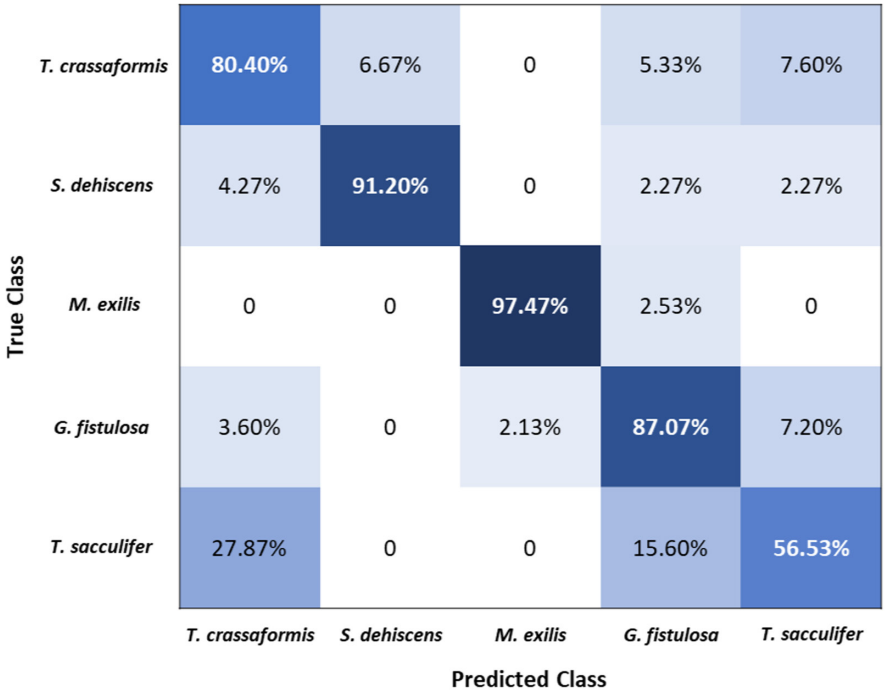

FIGURE 7 The mean confusion matrix of the classification performance from the 50 randomized partitions using the features extracted by 3DKMI. It shows that the samples belonging to the species *Menardella exilis* were nearly perfectly distinguished from the samples of other species, whereas the samples of the species *Trilobatus sacculifer* were not.

challenging conditions. Figure 5 summarizes the MSE between the 3D Krawtchouk moment invariants of the original object and the ones under random transformations. The variation in error arises from 50 random transformations, which were tested separately for both global and local settings. Despite the application of random transformations to the given object, both global and local features exhibit low error. However, it is worth noting that the error associated with local features is higher, with a median of $8e^{-4}$ compared to $6e^{-10}$ observed with global features. This difference is still related to the sensitivity of local features to scaling.

3.4 | Classification

We now show 3DKMI's power in exploiting the signatures of individual morphology to classify various species. All images were resized uniformly to a resolution of 128×128×128 and were described globally using the second order of the moments. Figure 6 shows the Krawtchouk moment morphospace filled by the randomly selected 250 individuals across the five species.

Subsequently, the formed dataset of these three features was partitioned into training and test sets, representing 70% and 30% of the data, respectively. The support vector machine (SVM) classifier, a classic method that aims to maximize the margin between classes and is robust to outliers, was employed to perform the classification task using the features extracted by 3DKMI. In detail, the SVM was fitted to the training set initially to learn the underlying patterns within the data, and then the trained SVM was applied to predict the classes of each sample in the test set. In experiments with 50 random partitions, classification accuracies on the test set ranged from 76% to 91%, with a median accuracy of 83%. Similar results were obtained with a k-nearest neighbours classifier ($k=5$). Figure 7 shows the mean confusion matrix generated from these 50 partitions, revealing the percentage of both correct and incorrect predictions for each class. By analysing the confusion matrix, it is evident that the classifier almost perfectly predicted samples belonging to the species *M. exilis*, and yet exhibited some confusion when classifying samples from the species *T. sacculifer*. *G. fistulosa* evolved directly from the *T. sacculifer* group of morphospecies in the late Pliocene (King et al., 2020); these two species differ through the existence of elongated protuberances in the final life stages. While some confusion in classification is unsurprising given that *T. sacculifer* and *T. crassaformis* share similar aspect ratios and our random sampling regime did not prioritize the most obvious protuberances, a major source of this confusion is the preservation (or not) of internal structures (Figure 3).

4 | CONCLUSIONS

Morphometric characterizations of individuals have emerged as a crucial tool to quantitatively compare biological forms. Automatic techniques that can effectively derive repeatable features to represent individuals have thus assumed significant roles in ecology and evolution. In this study, we developed a MATLAB package, 3DKMI, designed for automated feature extraction from any 3D image. Our package relies on the utilization of 3D Krawtchouk moment invariants, which are well-established tools in computer vision as they are renowned for being repeatable and transformation invariant, making them ideal for shape representation. Our experiments on this planktonic foraminifera dataset suggest that 3D Krawtchouk moment invariants can differentiate among traditionally delimited

species (Figures 6, 7), and the performance would increase further by incorporating other traits, particularly size, which is a clear differentiator between *T. crassaformis* and *T. sacculifer*, when building morphospaces.

The 3DKMI package possesses many distinctive advantages. It requires only a consumer-grade laptop for extremely fast and reproducible implementation and is based on an interpretable mathematical framework that efficiently computes the geometric features of an object directly from the input of a single image without the need to spend time and resources training a deep learning model on a dataset. The extracted features are repeatable and independent of observer bias. 3DKMI is also part of a growing effort to develop new methods that eliminate the reliance on landmarks and homology, contributing new translation, scaling, and rotation invariant tools for fast and repeatable extraction of phenotypic data from 3D volumetric images. We expect that 3DKMI will greatly expand the scope of morphometric research by enabling rapid comparisons of continually more disparate shapes.

AUTHOR CONTRIBUTIONS

Huahua Lin, Wenshu Zhang, Mark Nixon, Xiaohao Cai and Thomas H. G. Ezard conceived the ideas and designed the methodology; Anieke Brombacher, Alex Searle-Barnes and James M. Mulqueeney collected the data; Huahua Lin and Wenshu Zhang analysed the data; Huahua Lin led the writing of the manuscript, which was iterated with Xiaohao Cai and Thomas H. G. Ezard. All authors contributed critically to the drafts and gave final approval for publication.

FUNDING INFORMATION

This work was funded by the Natural Environment Research Council Award NE/P019269/1.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14388.

DATA AVAILABILITY STATEMENT

Data available from the Dryad Digital Repository https://doi.org/10.5061/dryad.66t1g1k93 (Lin et al., 2024).

ORCID

Huahua Lin https://orcid.org/0000-0002-1600-6293
Wenshu Zhang https://orcid.org/0000-0003-0776-7709
James M. Mulqueeney https://orcid.org/0000-0003-3502-745X
Anieke Brombacher https://orcid.org/0000-0003-2310-047X
Alex Searle-Barnes https://orcid.org/0000-0003-0389-7717
Mark Nixon https://orcid.org/0000-0002-9174-5934
Xiaohao Cai https://orcid.org/0000-0003-0924-2834
Thomas H. G. Ezard https://orcid.org/0000-0001-8305-6605

REFERENCES

Adam, A. (2023). *Mesh voxelisation*. https://www.mathworks.com/matlabcentral/fileexchange/27390-mesh-voxelisation. MATLAB Central File Exchange.

Alsmadi, M. K., Omar, K. B., & Noah, S. A. (2011). Fish classification based on robust features extraction from color signature using back-propagation classifier. *Journal of Computer Science*, *7*(1), 52–58.

Alsmadi, M. K., Omar, K. B., Noah, S. A., & Almarashdeh, I. (2010). Fish recognition based on robust features extraction from size and shape measurements using neural network. *Journal of Computer Science*, *6*(10), 1088–1094.

Bardua, C., Felice, R. N., Watanabe, A., Fabre, A.-C., & Goswami, A. (2019). A practical guide to sliding and surface semilandmarks in morphometric analyses. *Integrative Organismal Biology*, *1*(1), obz016.

Benouini, R., Batioua, I., Zenkouar, K., Najah, S., & Qjidaa, H. (2018). Efficient 3D object classification by using direct krawtchouk moment invariants. *Multimedia Tools and Applications*, *77*, 27517–27542.

Brombacher, A., Wilson, P. A., Bailey, I., & Ezard, T. H. (2021). The dynamics of diachronous extinction associated with climatic deterioration near the neogene/quaternary boundary. *Paleoceanography and Paleoclimatology*, *36*(6), e2020PA004205.

Collyer, M. L., Sekora, D. J., & Adams, D. C. (2015). A method for analysis of phenotypic change for phenotypes described by high-dimensional data. *Heredity*, *115*(4), 357–365.

Cooney, C. R., Bright, J. A., Capp, E. J., Chira, A. M., Hughes, E. C., Moody, C. J., Nouri, L. O., Varley, Z. K., & Thomas, G. H. (2017). Mega-evolutionary dynamics of the adaptive radiation of birds. *Nature*, *542*(7641), 344–347.

Cooney, C. R., Varley, Z. K., Nouri, L. O., Moody, C. J., Jardine, M. D., & Thomas, G. H. (2019). Sexual selection predicts the rate and direction of colour divergence in a large avian radiation. *Nature Communications*, *10*(1), 1773.

Devine, J., Aponte, J. D., Katz, D. C., Liu, W., Vercio, L. D. L., Forkert, N. D., Marcucio, R., Percival, C. J., & Hallgrímsson, B. (2020). A registration and deep learning approach to automated landmark detection for geometric morphometrics. *Evolutionary Biology*, *47*(3), 246–259.

Fabre, A.-C., Goswami, A., Peigné, S., & Cornette, R. (2014). Morphological integration in the forelimb of musteloid carnivorans. *Journal of Anatomy*, *225*(1), 19–30.

Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., & Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, *186*, 64–87.

Felice, R. N., Randau, M., & Goswami, A. (2018). A fly in a tube: Macroevolutionary expectations for integrated phenotypes. *Evolution*, *72*(12), 2580–2594.

Goswami, A., Smaers, J. B., Soligo, C., & Polly, P. D. (2014). The macroevolutionary consequences of phenotypic integration: From development to deep time. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *369*(1649), 20130254.

Hsiang, A. Y., Brombacher, A., Rillo, M. C., Mleneck-Vautravers, M. J., Conn, S., Lordsmith, S., Jentzen, A., Henehan, M. J., Metcalfe, B., Fenton, I. S., Wade, B. S., Fox, L., Meilland, J., Davis, C. V., Baranowski, U., Groeneveld, J., Edgar, K. M., Movellan, A., Aze, T., … Hull, P. M. (2019). Endless forams: >34,000 modern planktonic foraminiferal images for taxonomic training and automated species recognition using convolutional neural networks. *Paleoceanography and Paleoclimatology*, *34*(7), 1157–1177.

Hu, M.-K. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, *8*(2), 179–187.

King, D. J., Wade, B. S., Liska, R. D., & Miller, C. G. (2020). A review of the importance of the caribbean region in oligo-miocene low latitude

planktonic foraminiferal biostratigraphy and the implications for modern biogeochronological schemes. *Earth-Science Reviews*, *202*, 102968.

Lam, A. R., & Leckie, R. M. (2020a). Late neogene and quaternary diversity and taxonomy of subtropical to temperate planktic foraminifera across the kuroshio current extension, northwest pacific ocean. *Micropaleontology*, *66*(3), 177–268.

Lam, A. R., & Leckie, R. M. (2020b). Subtropical to temperate late neogene to quaternary planktic foraminiferal biostratigraphy across the kuroshio current extension, shatsky rise, northwest pacific ocean. *PLoS One*, *15*(7), e0234351.

Lin, H., Zhang, W., Mulqueeney, J. M., Brombacher, A., Searle-Barnes, A., Nixon, M., Cai, X., & Ezard, T. H. G. (2024). Data from: 3DKMI: A MATLAB package to generate shape signatures from Krawtchouk moments and an application to species delimitation in planktonic foraminifera. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.66t1g1k93

Mukundan, R., Ong, S., & Lee, P. A. (2001). Image analysis by tchebichef moments. *IEEE Transactions on Image Processing*, *10*(9), 1357–1364.

Mulqueeney, J. M., Ezard, T. H., & Goswami, A. (2024). Assessing the application of landmark-free morphometrics to macroevolutionary analyses. *bioRxiv*. 2024-04.

Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica*, *11*(285–296), 23–27.

Pearson, P. N., & Ezard, T. H. (2014). Evolution and speciation in the eocene planktonic foraminifer turborotalia. *Paleobiology*, *40*(1), 130–143.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-based species delimitation for the dna taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595–609.

Poole, C. R., & Wade, B. S. (2019). Systematic taxonomy of the *Trilobatus sacculifer* plexus and descendant *Globigerinoidesella fistulosa* (planktonic foraminifera). *Journal of Systematic Palaeontology*, *17*(23), 1989–2030.

Porto, A., Rolfe, S., & Maga, A. M. (2021). Alpaca: A fast and accurate computer vision approach for automated landmarking of three-dimensional biological structures. *Methods in Ecology and Evolution*, *12*(11), 2129–2144.

Porto, A., & Voje, K. L. (2020). Ml-morph: A fast, accurate and general approach for automated detection and landmarking of biological structures in images. *Methods in Ecology and Evolution*, *11*(4), 500–512.

Rolland, J., Henao-Diaz, L. F., Doebeli, M., Germain, R., Harmon, L. J., Knowles, L. L., Liow, L. H., Mank, J. E., Machac, A., Otto, S. P., Pennell, M., Salamin, N., Silvestro, D., Sugawara, M., Uyeda, J., Wagner, C. E., & Schluter, D. (2023). Conceptual and empirical bridges between micro- and macroevolution. *Nature Ecology & Evolution*, *7*, 1181–1193.

Siddiqi, K., Zhang, J., Macrini, D., Shokoufandeh, A., Bouix, S., & Dickinson, S. (2008). Retrieving articulated 3-D models using medial surfaces. *Machine Vision and Applications*, *19*, 261–275.

Sit, A., Shin, W.-H., & Kihara, D. (2019). Three-dimensional krawtchouk descriptors for protein local surface shape comparison. *Pattern Recognition*, *93*, 534–545.

Spampinato, C., Giordano, D., Di Salvo, R., Chen-Burger, Y.-H. J., Fisher, R. B., & Nadarajan, G. (2010). Automatic fish classification for underwater species behavior understanding. *Proceedings of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams*. New York, NY, USA, pp. 45–50.

Talavera, G., Dincă, V., & Vila, R. (2013). Factors affecting species delimitations with the gmyc model: Insights from a butterfly survey. *Methods in Ecology and Evolution*, *4*(12), 1101–1110.

Teague, M. R. (1980). Image analysis via the general theory of moments. *Journal of the Optical Society of America*, *70*(8), 920–930.

Viacava, P., Blomberg, S. P., & Weisbecker, V. (2023). The relative performance of geometric morphometrics and linear-based methods in the taxonomic resolution of a mammalian species complex. *Ecology and Evolution*, *13*(3), e9698.

Wäldchen, J., & Mäder, P. (2018). Machine learning for image based species identification. *Methods in Ecology and Evolution*, *9*(11), 2216–2225.

Woodhouse, A., Procter, F. A., Jackson, S. L., Jamieson, R. A., Newton, R. J., Sexton, P. F., & Aze, T. (2023). Paleoecology and evolutionary response of planktonic foraminifera to the mid-pliocene warm period and plio-pleistocene bipolar ice sheet expansion. *Biogeosciences*, *20*(1), 121–139.

Yap, P.-T., Paramesran, R., & Ong, S.-H. (2003). Image analysis by krawtchouk moments. *IEEE Transactions on Image Processing*, *12*(11), 1367–1377.

Zelditch, M. L., Swiderski, D. L., & Sheets, H. D. (2012). *Geometric morphometrics for biologists: A primer*. Academic Press.

How to cite this article: Lin, H., Zhang, W., Mulqueeney, J. M., Brombacher, A., Searle-Barnes, A., Nixon, M., Cai, X., & Ezard, T. H. G. (2024). 3DKMI: A MATLAB package to generate shape signatures from Krawtchouk moments and an application to species delimitation in planktonic foraminifera. *Methods in Ecology and Evolution*, *15*, 1940–1948. https://doi.org/10.1111/2041-210X.14388