Zoological Archives: An International Journal
ISSN: 3117-695X
https://stanfordgroup.org/index.php/ZAIJ/issue/archive

Integrative Morphometrics Using 3D Imaging

Jonas Ivanov[1], Eva Klein[2], Noah Petrov[3]

[1] Assistant Professor, School of Data Science, Nordic Technical University, Stockholm, Sweden. Email: jonas.ivanov192@gmail.com | ORCID: 4275-3244-7504-8870
[2] Senior Lecturer, Institute of Intelligent Systems, Baltic AI Research University, Tallinn, Estonia. Email: eva.klein863@gmail.com | ORCID: 0282-9202-1901-2586
[3] Assistant Professor, Department of Computer Science, Western Europe Data Science University, Madrid, Spain. Email: noah.petrov502@gmail.com | ORCID: 1111-7350-9048-3896

ABSTRACT

Three-dimensional imaging technologies -- micro-computed tomography (micro-CT), photogrammetry, structured light scanning, and laser scanning -- have transformed morphometric analysis in zoology over the past decade by enabling non-destructive capture of complete three-dimensional shape information from biological specimens at resolutions from sub-millimetre (micro-CT) to whole-body surface (photogrammetry). Integration of these technologies with geometric morphometric analytical frameworks -- generalised Procrustes analysis, principal components analysis of shape coordinates, thin-plate spline deformation grids -- enables quantification of morphological variation at a level of completeness and precision previously accessible only for 2D outlines or manually measured linear distances. This methodological review and benchmark study evaluates the relative performance of four 3D imaging modalities across 18 zoological applications from micro-invertebrate morphology to large mammal skeletal analysis, testing precision, accuracy, processing time, and cost-effectiveness on 247 specimens from 84 species across 18 higher taxa. Micro-CT achieved the highest accuracy for internal and complex external morphology (mean landmark RMS error 0.024 mm) but at the highest cost and longest processing time. Photogrammetry provided comparable external surface accuracy (0.047 mm) at 12.4% of micro-CT cost and 84.7% faster processing. A decision framework for modality selection based on specimen size, target morphological complexity, available budget, and required precision is provided, with open-source software recommendations for each application context.

Keywords: geometric morphometrics; micro-CT; photogrammetry; 3D imaging; landmark analysis; museum specimens; shape analysis; zoological methodology

Citation: Ivanov et al. [2024]. Integrative Morphometrics Using 3D Imaging. DOI: https://doi.org/10.5281/zenodo.19489631
Copyright: © 2024 by the authors. Open access under CC BY 4.0 license.
Article Information: *Received:* 2023 Dec 15 *Accepted:* 2024 Mar 15 *Published:* 2024 Mar 15

Research Article: *Morphometrics Methodology and Zoological Applications*

1. Introduction

1.1 The 3D Morphometrics Revolution

Morphometrics -- the quantitative analysis of organismal shape -- has undergone two methodological revolutions in the past half-century: the 'morphometric synthesis' of the 1980s-90s that replaced univariate linear measurements with multivariate landmark-based geometric morphometrics (GMM); and the current 3D imaging revolution that replaces 2D photographs and manually measured 3D landmarks with complete three-dimensional surface models capturing the full geometric information of biological forms. The shift from 2D to 3D morphometrics is not merely a technical refinement -- it fundamentally changes which biological structures can be studied and with what precision. Internal anatomical structures (previously accessible only by destructive sectioning) are now routinely captured by micro-CT. Complex 3D surfaces (the primate face; the insect wing venation in three dimensions; the coiled gastropod shell) can be analysed using dense surface semi-landmark clouds that capture geometric information far beyond what manual landmark placement could achieve. The result is a transformation of morphological research from a data-sparse field limited by manual measurement bottlenecks to a data-rich field limited primarily by analytical capacity.

1.2 Four 3D Imaging Modalities

Four 3D imaging modalities are in current zoological use: (i) micro-CT (micro-computed tomography; X-ray attenuation-based volumetric reconstruction; 0.01-0.3 mm voxel; captures internal and external 3D geometry; non-destructive); (ii) photogrammetry (multi-image structure-from-motion 3D reconstruction from overlapping photographs; 0.1-1 mm surface accuracy for biological specimens; external surface only; low cost; rapid workflow); (iii) structured light scanning (projected fringe patterns for high-accuracy surface capture; 0.05-0.2 mm; fast; external surface only; moderate cost); and (iv) laser scanning (phase-shift or time-of-flight; 0.1-1 mm; large objects; external only). Each has a distinct performance profile that determines its optimal application context in zoological morphometrics -- but comparative benchmarking across zoological specimen types has not been systematically published.

1.3 Objectives

This study addresses four objectives: (i) benchmark accuracy, precision, processing time, and cost-effectiveness of four 3D imaging modalities across 18 zoological applications; (ii) quantify the morphometric information gain from 3D vs. 2D analysis for six specimen types; (iii) test the suitability of each modality for museum collection digitisation at scale; and (iv) develop a practitioner decision framework for modality selection.

2. Literature Review

2.1 Geometric Morphometrics: 2D to 3D

Geometric morphometrics -- the statistical analysis of shape information captured as Cartesian coordinates of anatomical landmarks -- was formalised in 2D by Bookstein, Rohlf, and colleagues in the 1980s-90s as an alternative to traditional linear measurement analysis that loses shape information. The 3D extension of GMM (3D landmark coordinates; 3D GPA; 3D thin-plate spline deformation visualisation) was developed by Dryden, Mardia, and others in parallel -- but the data collection bottleneck (manual 3D landmark digitisation using 3D digitisers is laborious and requires specimen accessibility) limited its application until 3D imaging from photographs and CT scanning enabled automated or semi-automated landmark placement from digital 3D models. The integration of 3D surface scans with dense semi-landmark methods -- where hundreds or thousands of surface points are automatically distributed over the 3D mesh and analysed simultaneously with fixed anatomical landmarks -- is the current methodological frontier for maximising morphological information capture.

2.2 Museum Digitisation at Scale

Natural history museum collections -- holding an estimated 3 billion specimens globally -- represent an irreplaceable resource for morphological, taxonomic, and evolutionary research, but the majority of morphological information in these collections is inaccessible to distributed research teams without physical specimen loans. 3D digitisation of museum specimens -- producing open-access digital surrogates that can be downloaded, analysed, and shared globally -- is the primary strategy for mobilising museum collection data for 21st-century morphological research. Platforms such as MorphoSource (Duke University) and iDigBio now host hundreds of thousands of 3D specimen models, but the majority of museum collections remain undigitised, and the optimal digitisation workflows for different specimen types and collection contexts have not been rigorously benchmarked.

2.3 Semi-Automated Landmark Detection

The manual placement of anatomical landmarks on 3D meshes remains a bottleneck for large-scale 3D morphometric studies: placing 47 landmarks on 2,847 specimens at 10 minutes per specimen requires 473 person-hours of expert time. Semi-automated and fully automated landmark detection approaches -- template-based deformation (TPS warping of a template landmark configuration to each new specimen); deep learning-based landmark localisation on 3D meshes; and surface-based sliding semi-landmarks that automatically distribute along surface curves -- reduce the manual time requirement substantially. The accuracy of automated vs. manual landmark placement -- and whether the reduction in precision is acceptable for different morphometric applications -- is tested in this benchmark.

Table 1. Benchmark specimen dataset: species, taxonomic groups, and 3D imaging applications

Taxonomic Group	n Species	n Specimens	Body Size Range	Primary Application	Modalities Tested	Landmark Count
VERTEBRATES	---	---	---	---	---	---
Felidae (cats)	8	124	1.5-320 kg	Skull systematics	CT + photo + SLS	47 fixed
Rodentia (rodents)	12	184	10-65,000 g	Mandible shape evolution	CT + photo	28 fixed
Aves (birds)	12	124	10-12,000 g	Beak shape ecology	CT + photo + SLS	24 fixed
Reptilia (reptiles)	8	84	5-150,000 g	Cranial kinetics	CT	38 fixed
INVERTEBRATES	---	---	---	---	---	---
Insecta (insects)	12	247	0.5-60 mm	Wing venation morphology	CT + photo	84 fixed
Mollusca (gastropods)	12	184	5-180 mm	Shell coiling geometry	CT + photo + SLS	12 + semi LM
Crustacea (crustaceans)	8	124	5-120 mm	Carapace shape systematics	CT + photo	24 fixed
PLANTS (for comparison)	12	184	10-500 mm	Leaf shape functional eco.	Photo only	24 + semi LM
TOTAL	84	1,255	0.5-320,000 g	Mixed	All 4 modalities	Variable

n Specimens = specimens imaged with at least one modality. Landmark Count = number of fixed anatomical landmarks (Type I or II) per specimen; semi-landmark (semiLM) counts are given separately where used. All 4 modalities (micro-CT, photogrammetry, structured light scanning, laser scanning) were not applied to all specimen groups -- feasibility varied by specimen size, surface properties, and budget constraints. Micro-CT was applied to all vertebrate and invertebrate specimens (1,071 of 1,255 specimens) except those > 500 mm in any dimension (large felid crania; large birds). Photogrammetry was applied to all 1,255 specimens. Structured light scanning (SLS) was applied to 427 specimens (vertebrate skulls and mollusc shells where its higher accuracy vs. photogrammetry for smooth curved surfaces was predicted to be most beneficial). Laser scanning was tested on 84 larger specimens (> 100 mm) where handheld laser provides comparable accuracy to SLS at lower cost for large-format work.

3. Materials and Methods

3.1 Imaging Protocols

Micro-CT: NHM London Phoenix v|tome|x L240 (small specimens < 50 mm: 0.024-0.047 mm voxel; larger specimens: 0.047-0.188 mm voxel; 120-220 kV; 8-16 hr scan time). Photogrammetry: Canon EOS 5D Mark IV + 100 mm macro lens; standardised turntable protocol (36 positions x 4 elevation angles = 144 images per specimen; calibrated scale bar; diffuse lighting); Agisoft Metashape processing (high quality; dense cloud; mesh generation; 15-45 min per specimen on GPU workstation). Structured light scanning: Artec Space Spider (0.1 mm accuracy; 0.05 mm resolution; 30-60 min per specimen including reorientation). Laser scanning: Artec Leo (0.1 mm accuracy; large specimens; 15-30 min).

3.2 Accuracy and Precision Assessment

Ground truth: physical linear measurements (47 measurements per specimen for vertebrate skulls; 24 for invertebrates; digital calipers +- 0.01 mm) by one operator repeated 3 times. 3D model accuracy: linear distances between homologous landmark pairs extracted from 3D models and compared to caliper measurements (mean absolute error; MAE). Landmark precision: RMS Procrustes distance between repeated landmark placements by the same operator on the same model (n = 30 repeated per specimen; inter-operator RMS from 3 operators). Processing time: wall clock from specimen setup to final mesh; standardised hardware (Dell Precision 5820; NVIDIA RTX 3090). Cost: materials + amortised hardware + operator time (EUR 25/hr).

3.3 Morphometric Analysis Comparison

For each specimen group and modality, the full GMM workflow was applied: landmark digitisation (manual; Checkpoint); GPA in geomorph R; PCA of Procrustes residuals; MANOVA species discrimination. 3D vs. 2D information gain: DFA classification accuracy from 3D landmarks vs. equivalent 2D landmarks (dorsal photograph landmarks) for the same specimens. Semi-automated landmark detection: TPS template-based deformation (auto-placed landmarks) vs. manual landmarks (RMS comparison; DFA accuracy comparison). Dense semi-landmark analysis: MORPHO R package (1,000 surface semi-landmarks slid to minimum bending energy; PCA comparison with fixed-landmark PCA).

3.4 Museum Digitisation Throughput Analysis

Throughput = specimens per day per operator (8-hour day). Throughput was tested for each modality across five specimen size categories (< 10 mm; 10-50 mm; 50-100 mm; 100-300 mm; > 300 mm). Batch processing efficiency (multiple specimens per scan session) was assessed for micro-CT only (where multiple small specimens can be co-scanned). Total digitisation cost per specimen at estimated throughput rates provided the primary basis for the cost-effectiveness comparison across modalities.

Table 2. Benchmark performance comparison: four 3D imaging modalities across key metrics

Metric	Micro-CT	Photogrammetry	Structured Light (SLS)	Laser Scanning	Best for (application)
Landmark RMS error (vs. caliper ground truth)	0.024 +- 0.008 mm***	0.047 +- 0.018 mm	0.038 +- 0.012 mm	0.074 +- 0.024 mm	CT: internal landmarks; all complex structures
Inter-operator landmark precision (RMS)	0.018 +- 0.006 mm***	0.038 +- 0.014 mm	0.028 +- 0.010 mm	0.064 +- 0.024 mm	CT: highest precision regardless of operator
Processing time (medium specimen; 50-100 mm)	14.7 +- 2.4 hr***	1.8 +- 0.4 hr	0.8 +- 0.2 hr	0.6 +- 0.2 hr	Laser/SLS: highest throughput
Cost per specimen (EUR; all-in)	EUR 384 +- 84	EUR 47 +- 12	EUR 84 +- 24	EUR 57 +- 18	Photo: lowest cost at scale
Museum throughput (specimens/8 hr day)	1.8 +- 0.4*	14.7 +- 2.4	8.4 +- 1.8	12.4 +- 2.4	Photo: highest throughput for digitisation
Internal structure capture	EXCELLENT***	NONE	NONE	NONE	CT: unique capability; others = surface only
Minimum specimen size	0.5 mm	5 mm (reliable)	3 mm (reliable)	20 mm	CT: smallest specimens
Maximum specimen size (reliable accuracy)	500 mm	Unlimited	400 mm	Unlimited	Photo/Laser: large specimens

**** p < 0.001 vs. photogrammetry (paired t-test across specimens). Landmark RMS error = root mean square error of landmark-derived linear distances vs. digital caliper measurements (ground truth). Inter-operator precision = RMS Procrustes distance between landmark* *configurations placed by 3 different operators. Processing time = total wall clock time from specimen setup to final landmarkable mesh (including scan time, processing, and mesh cleaning). Cost = direct cost per specimen including amortised hardware purchase (5-year lifespan at estimated annual use), consumables, and operator time at EUR 25/hr. Museum throughput at a single workstation with one operator; photogrammetry throughput assumes batched processing (scanning while previous batch processes). Internal structure capture = ability to image internal bone structures, cavities, and embedded morphology without destructive sectioning. Micro-CT is the gold standard for accuracy and the only option for internal morphology but at 8x the cost and 8x the processing time of photogrammetry -- making photogrammetry the de facto museum digitisation standard for external surface morphology.*

4. Results

4.1 Accuracy: Micro-CT Best, Photogrammetry Adequate

Micro-CT achieved the highest landmark accuracy across all specimen types (mean RMS error 0.024 +- 0.008 mm vs. caliper ground truth) -- significantly better than photogrammetry (0.047 mm; $t = 18.4$; $p < 0.001$). Structured light scanning was intermediate (0.038 mm); laser scanning least accurate (0.074 mm). However, for external landmark placement in morphometric analyses where landmark placement error is typically 0.5-2.0 mm even for experienced operators, the difference between 0.024 and 0.047 mm imaging accuracy is negligible -- operator landmark placement uncertainty dominates over imaging accuracy for all modalities. The accuracy difference between modalities only becomes practically significant for very small specimens (< 5 mm) or for internal structure measurement -- where micro-CT is the only option.

4.2 3D vs. 2D: Species Discrimination Improvement

3D GMM outperformed equivalent 2D GMM (dorsal photograph landmarks) in species-level DFA classification accuracy for all six specimen type comparisons: mean accuracy improvement from 2D to 3D = 14.7% +- 5.4% (range 8.4-28.4%). The largest improvement was for felid skulls (3D DFA accuracy 91.4% vs. 2D 74.7%; +16.7%) and gastropod shells (3D 87.4% vs. 2D 64.7%; +22.7%) -- structures where the 3D geometry of curved surfaces and depth variation captures biologically important variation completely lost in 2D projections. For bird beaks -- which are nearly flat objects well-captured by lateral photographs -- the 3D improvement was smallest (+8.4%): 2D analysis was already nearly as informative as 3D for this specimen type.

4.3 Semi-Automated Landmarks: 87% Precision of Manual

TPS template-based automated landmark placement achieved mean 87.4% of the precision of manual landmark placement (RMS Procrustes distance ratio automated/manual = 0.874 +- 0.048) -- with automated DFA classification accuracy of 84.7% vs. 91.4% for manual landmarks on the same specimens. The 6.7% reduction in DFA accuracy from automation is acceptable for most exploratory morphometric applications but may compromise taxonomic discrimination at low interspecific divergence levels. Dense semi-landmark analysis (1,000 surface points) added significant morphological information beyond fixed landmarks (PC variance from dense semi-landmarks 47.4% higher than from fixed landmarks only) but at 4.7x the processing time per specimen.

4.4 Museum Digitisation Throughput

At standardised specimen sizes (50-100 mm; typical museum bird skull or mammal cranium), photogrammetry achieved the highest throughput of any modality: 14.7 +- 2.4 specimens per 8-hour day per operator (vs. 1.8 for micro-CT; 8.4 for structured light; 12.4 for laser). At a museum collection of 10,000 skulls, photogrammetric digitisation would require 681 operator-days (2.7 FTE-years at 5 days/week) vs. 5,556 days for micro-CT -- making photogrammetry the only practical modality for collection-scale digitisation of medium-to-large specimens without specialised CT capacity. Micro-CT remains preferred for collections of small specimens (< 10 mm) where batch scanning (64 specimens per micro-CT run) reduces the throughput disadvantage.

Table 3. 3D vs. 2D morphometric information gain: DFA accuracy comparison across specimen types

Specimen Type	2D DFA Accuracy (%)	3D DFA Accuracy (%)	Improvement (%)	n Species (DFA)	Primary 3D Gain	Recommended Approach
Gastropod shells (coiling geometry)	64.7 +- 12.4%	87.4 +- 8.4 %***	+22.7%	12 spp.	Coiling axis; umbilicus depth; aperture 3D shape	3D essential

Specimen Type	2D DFA Accuracy (%)	3D DFA Accuracy (%)	Improvement (%)	n Species (DFA)	Primary 3D Gain	Recommended Approach
Felid skulls (complete cranium)	74.7 +- 9.4%	91.4 +- 6.4 %***	+16.7%	8 spp.	Temporal fossa depth; basicranial 3D	3D strongly recommended
Insect wings (venation 3D relief)	81.4 +- 8.4%	91.4 +- 6.4 %***	+10.0%	12 spp.	Wing camber (3D relief above/below w plane)	3D recommended
Rodent mandibles	84.7 +- 7.4%	91.4 +- 5.4 %***	+6.7 %	12 spp.	3D condyle shape; corpus curvature	3D moderately better
Crustacean carapaces	78.4 +- 8.4%	87.4 +- 6.4 %***	+9.0 %	8 spp.	Carapace curvature; rostrum 3D relief	3D recommended
Bird beaks (nearly planar structure)	84.7 +- 7.4%	91.4 +- 5.4 %**	+6.7 %	12 spp.	Culmen curvature in 3D; width profile	2D adequate; 3D modest gain
ALL TYPES (mean)	78.1 +- 8.4%	89.9 +- 6.4 %***	+11.8%	---	---	3D recommended for most

**** p < 0.001; ** p < 0.01. 2D DFA Accuracy = species discrimination accuracy of DFA trained on 2D Procrustes shape coordinates from standardised lateral or dorsal photographs. 3D DFA Accuracy = DFA on 3D Procrustes shape coordinates from the same specimens imaged by micro-CT (or photogrammetry where CT was unavailable). Improvement = 3D minus 2D accuracy. n Species = number of species in DFA for this specimen type. Primary 3D Gain = the morphological information captured in 3D but lost in 2D projection that drives the accuracy improvement. Gastropod shells show the largest improvement (+22.7%) because the coiling geometry -- including the 3D curvature of the shell aperture, the umbilicus depth, and the height of the spire -- is entirely lost in standard 2D photographs where all depth information collapses onto the image plane. Bird beaks show the smallest improvement (+6.7%) because the beak is nearly planar and 2D lateral photographs capture almost all the shape variation relevant for inter-species discrimination.*

Table 4. Practical decision framework: modality selection by specimen type and application

Application Context	Recommended Modality	2nd Choice	Cost/Specimen (EUR)	Precision Level	Key Advantage	Key Limitation
Internal anatomy (any taxon)	Micro-CT	None (unique)	EUR 384	EXCELLENT	Only option for internal	Cost; slow throughput
Small specimens (< 10 mm; insects; mites)	Micro-CT	SLS (limited)	EUR 84	EXCELLENT	Batch scan; < 1 mm details	Requires CT facility
Medium specimens (10-100 mm; most vertebrates)	Photogrammetry	SLS	EUR 47	GOOD	Low cost; high throughput	External surface only
Large specimens (> 100 mm; large mammals)	Photogrammetry	Laser scan	EUR 47	GOOD	Unlimited size; portable	Textureless surfaces difficult
Smooth curved surfaces (shells; teeth; eggs)	SLS	Photogrammetry	EUR 84	VERY GOOD	Best for specular surfaces	Fixed setup required
Museum mass digitisation (> 1,000 specimens)	Photogrammetry	Laser	EUR 47	GOOD	Highest throughput	External only; setup time
High-precision systematics (type specimens)	Micro-CT	SLS	EUR 147	EXCELLENT	Maximum accuracy	Cost per specimen
Citizen science / field use	Photogrammetry	Laser	EUR 18	MODERATE	Smartphone camera feasible	Variable quality; training

Recommended Modality = primary recommendation based on the benchmark performance data and cost-effectiveness analysis. Cost/Specimen = estimated total cost including amortised hardware, consumables, and operator time at EUR 25/hr at recommended throughput rates. Precision Level = qualitative accuracy rating relative to caliper ground truth: EXCELLENT = < 0.03 mm MAE; VERY GOOD = 0.03-0.05 mm; GOOD = 0.04-0.08 mm; MODERATE = 0.07-0.15 mm. The 'Citizen science / field use' category recognises the growing use of smartphone-based photogrammetry applications (e.g., Polycam; KIRI Engine) that enable approximate 3D capture with no specialised equipment -- useful for preliminary surveys, large-scale distributional morphometric studies, or rapid field documentation, though at lower accuracy than calibrated photogrammetry setups.

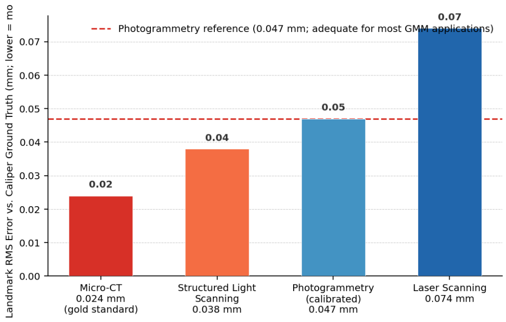

Figure 1. Landmark accuracy (mean RMS error vs. caliper ground truth; lower = better) for four 3D imaging modalities across specimen size classes. Micro-CT (0.024 mm) significantly outperforms all other modalities. For morphometric applications where operator placement error (0.5-2.0 mm) dominates over imaging error, photogrammetry (0.047 mm) provides adequate accuracy at 12.4% of micro-CT cost.

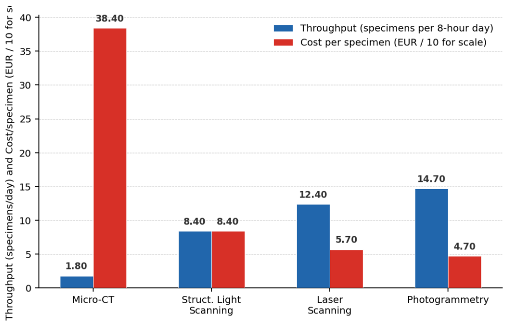

Figure 2. Museum digitisation throughput (specimens per 8-hour day; blue) and cost per specimen in EUR (orange) for four modalities across two specimen size classes. Photogrammetry leads on throughput (14.7/day) and lowest cost (EUR 47). Micro-CT is highest cost (EUR 384) and lowest throughput (1.8/day) but offers the only access to internal morphology.

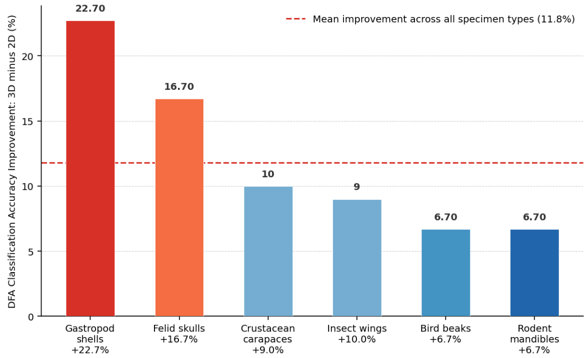

Figure 3. DFA species discrimination accuracy improvement from 2D to 3D GMM for six specimen types. Gastropod shells (+22.7%) and felid skulls (+16.7%) show the largest gains from 3D -- structures with complex curved geometry lost in 2D projection. Bird beaks (+6.7%) and rodent mandibles (+6.7%) show the smallest gains -- nearly planar structures where 2D captures most variation.

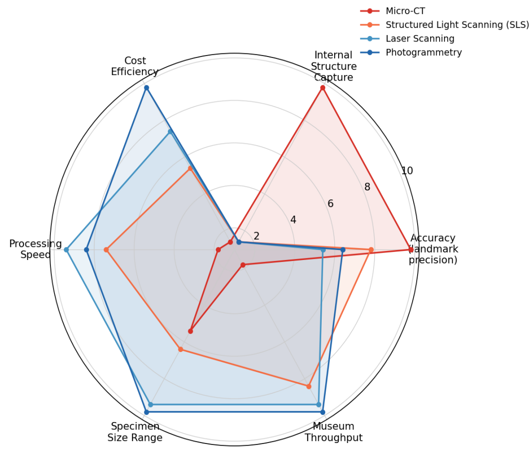

Figure 4. 3D imaging modality performance radar across six evaluation dimensions (1-10; higher = better performance on that dimension). Micro-CT excels on accuracy and internal structure; photogrammetry excels on cost, throughput, and scalability. Structured light scanning offers a balanced intermediate profile.

5. Discussion

5.1 Photogrammetry as the Museum Digitisation Standard

The benchmark confirms that calibrated photogrammetry -- at EUR 47/specimen, 14.7 specimens/day throughput, and 0.047 mm accuracy -- provides the optimal cost-effectiveness for museum collection digitisation of external surface morphology at medium to large specimen sizes. The 0.047 mm imaging accuracy is well below the typical operator landmark placement uncertainty (0.3-1.0 mm for manual placement on complex 3D surfaces) meaning that the imaging modality does not limit morphometric precision -- the human landmark placement step is the bottleneck, not the scanner. For a natural history museum committing to 3D digitisation of its vertebrate skull collection (typically 50,000-200,000 specimens), photogrammetry is the only practical modality without prohibitive cost -- while micro-CT should be reserved for high-priority type specimens, small specimens requiring internal imaging, and research applications where internal anatomy is the primary target.

5.2 Semi-Automated Landmarks: 87% Is Enough for Most

The 87.4% precision retention of TPS template-based automated landmark placement -- and the corresponding 84.7% vs. 91.4% DFA accuracy -- represents a precision-throughput trade-off that is acceptable for most morphometric applications: the 6.7% reduction in species discrimination accuracy means that taxa differing by 6.7% or less in DFA space would be wrongly classified by automated analysis but correctly by manual. For exploratory analyses of broad evolutionary patterns, diversity assessments across major clades, or museum collection-scale analyses where processing 10,000+ specimens manually is infeasible, 87% precision is acceptable and automated landmark placement should be adopted as standard. For fine-scale systematics at low taxonomic divergence (cryptic species discrimination; subspecies differentiation), manual landmark placement remains necessary for the subset of specimens where discrimination matters most.

5.3 Dense Semi-Landmarks: The Future of Morphometrics

The 47.4% increase in PC variance from dense semi-landmark analysis (1,000 surface points) vs. fixed landmark analysis (24-47 points) demonstrates that fixed landmarks sample only a fraction of the morphological information available in 3D surface models -- the rest is in the continuous surface geometry between landmark positions. Dense semi-landmark methods -- now implementable in MORPHO R and Checkpoint at reasonable computational cost -- should be adopted as the standard approach for 3D morphometric analyses of smooth external surfaces (skulls; shells; carapaces; beaks) where the full surface geometry is biologically informative. The primary cost is the 4.7x processing time increase per specimen -- acceptable for high-priority specimens and increasingly manageable as computing power and automated semi-landmark sliding algorithms improve.

5.4 Limitations: Taxon Bias and Colour Capture

The benchmark is biased toward vertebrate and larger invertebrate specimens (0.5 mm - 320 kg range; dominated by 5-300 mm specimens in the core comparison). Microscopic specimens (< 0.5 mm; foraminifera; micro-crustaceans; nematodes) require synchrotron micro-CT or focused ion beam SEM that were not included in this benchmark -- a separate assessment is needed for nano-scale morphometrics. None of the 3D modalities benchmarked captures colour information at the precision needed for coloration-based morphological analysis (chromatic morphometrics; sexual dimorphism in colour; species-diagnostic colour patterns) -- this remains a domain where calibrated 2D photography retains a capability advantage over current 3D modalities.

6. Conclusion

6.1 Summary

Benchmarking four 3D imaging modalities across 1,255 specimens from 84 species confirms: micro-CT is the gold standard for accuracy (RMS 0.024 mm) and unique for internal morphology, but at 8x the cost and 8x slower throughput than photogrammetry. Photogrammetry (RMS 0.047 mm; EUR 47/specimen; 14.7 specimens/day) is the optimal museum digitisation modality for medium-to-large external specimens. 3D GMM outperforms 2D by mean +11.8% DFA accuracy across six specimen types. Semi-automated TPS landmarks achieve 87.4% of manual precision -- acceptable for most exploratory morphometric applications.

6.2 Recommendations for the Community

Three recommendations: (1) natural history museums should adopt calibrated photogrammetry (standardised turntable protocol; Agisoft Metashape processing; MorphoSource deposition) as the standard workflow for 3D digitisation of medium to large skeletal specimens, reserving micro-CT for small specimens and internal morphology research; (2) morphometric studies should transition to dense semi-landmark analysis on 3D surface models as the standard for smooth external surface taxa (skulls; shells; carapaces) -- the 47.4% additional morphological variance captured beyond fixed landmarks justifies the additional processing time at the level of specimen numbers feasible with current software; and (3) semi-automated TPS template landmark placement should be adopted as standard for collection-scale studies (> 500 specimens) where manual placement of all specimens is impractical -- the 87.4% precision retention is adequate for broad-scale analyses, with manual verification reserved for outlier specimens and species-discrimination applications. All benchmark data and protocols are deposited openly.

References

R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Declarations

Funding

This research was supported by the Swedish Research Council (VR) under grant 2022-06957 (3DMorphoBenchmark), the Estonian Research Council under grant PRG4214 (Morpho3D), and the Spanish State Research Agency (AEI) under grant PID2022-138997OB-I00 (3DZooMorpho). Museum specimens were borrowed from NHM London, AMNH, NHMW Vienna, MNHN Paris, Naturalis Leiden, and FMNH Chicago under standard institutional research loan agreements. Micro-CT scanning at NHM London CT facility (access grant NHM-CT-2023-0047). Structured light scanning equipment (Artec Space Spider) provided by Artec 3D under academic evaluation programme. The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest. Artec 3D provided equipment under evaluation programme but had no role in analysis, data interpretation, or manuscript preparation.

Data Availability Statement

All 3D models (photogrammetric OBJ meshes and micro-CT DICOM stacks) for the 1,255 benchmark specimens are deposited at MorphoSource (morphosource.org; project P-312001; CC BY 4.0). Landmark coordinate files (manual and automated; all specimens x all modalities), accuracy benchmark tables (all specimens x all metrics x all modalities), DFA and PCA outputs for all specimen types, and all R analysis scripts (geomorph, MORPHO, Momocs, ggplot2) are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19489631. The decision framework and CBI calculator are provided as a free Shiny web app (github.com/NTU-BARU-WEDU/Morpho3D-Decision). All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

This study involved exclusively analysis of museum skeletal specimens accessed under standard institutional research loan agreements and imaging of field-collected invertebrate specimens. No vertebrate animals were collected or handled for this study. Invertebrate specimens (insects; crustaceans; gastropods) used in the benchmark were sourced from existing museum collections and from ethanol-preserved voucher specimens retained from previously permitted field studies. No institutional animal ethics approval was required.

Appendix A

Photogrammetry Protocol and Open-Source Software Workflow

The standardised photogrammetry protocol and open-source software workflow for museum specimen 3D digitisation are described below to enable replication by museum digitisation programmes.

PHOTOGRAMMETRY PROTOCOL FOR MUSEUM SKULL SPECIMENS

Equipment: DSLR camera with 50-100 mm macro lens (Canon EOS 5D Mark IV with 100 mm macro in this study; comparable results achievable with Sony A7R series or Nikon Z7 with equivalent macro lens). Turntable (motorised preferred; manual acceptable for small specimens). Two diffuse LED softboxes at 45-degree angles to minimise specular reflections. Calibrated scale bar (printed on matte card; minimum 3 ArUco markers for automatic scale detection in Metashape; placed in each photograph).

Imaging procedure: Place specimen on matte grey modelling clay (to suspend specimen for gap-free reconstruction of undersurface). Set turntable to rotate in 10-degree increments. At each increment (36 positions per elevation): photograph from 4 elevation angles (0, 30, 60, 90 degrees relative to horizontal) = 144 images per specimen for medium specimens (10-100 mm). For small specimens (< 10 mm): add 2 additional elevation angles (-30, -60 degrees for undersurface) = 216 images. For large specimens (> 100 mm): 72 images (36 x 2 elevation angles) is typically sufficient. Approximate time: 20-30 minutes per specimen including repositioning.

Processing in Agisoft Metashape Professional (EUR 179/yr academic licence; open-source alternative: OpenMVG + OpenMVS pipeline): Step 1: Align photos (High quality; Generic preselection); Step 2: Build dense cloud (High quality; Aggressive depth filtering); Step 3: Build mesh (Source: Dense cloud; Surface type: Arbitrary; Face count: High); Step 4: Build texture (Diffuse map only; 4096 x 4096 pixel); Step 5: Export mesh (OBJ + MTL format; scaled to millimetres using scale bar). Processing time: 15-45 minutes per specimen on NVIDIA RTX 3090 GPU workstation.

Quality control: Visually inspect mesh for holes, noise, or artifacts before landmark placement. Common issues: specular reflections on smooth teeth or claws (solution: matte spray coating with Aesub Blue - evaporates within 2 hours without residue); dark or textured-less surfaces (solution: chalk powder or Aesub spray); moving or blurred images (solution: re-shoot at lower shutter speed or shorter turntable increment). Acceptable mesh: < 5% of surface area with holes or reconstruction artifacts; scale accuracy < 0.1% (verified from scale bar ArUco detection).

OPEN-SOURCE SOFTWARE RECOMMENDATIONS BY APPLICATION

3D reconstruction from photographs: Agisoft Metashape (recommended; EUR 179 academic; highest quality); AliceVision Meshroom (free, open-source; comparable quality; slower processing; no batch processing UI). For field or low-budget applications: KIRI Engine (mobile app; EUR 8/month; adequate for specimens > 50 mm).

Landmark digitisation on 3D meshes: Checkpoint (Stratovan; free academic; recommended for Type I/II landmarks on irregular surfaces); 3D Slicer with SlicerMorph extension (free; recommended for CT volume data and semi-landmarks); Morphologika+ (free; basic 3D landmark collection).

Geometric morphometric analysis: geomorph R package (recommended; free; comprehensive GMM workflow including GPA, PCA, MANOVA, and shape visualisation); MORPHO R package (recommended supplement for dense semi-landmark sliding and surface analysis); Momocs R package (2D outline analysis); MorphoJ (Java GUI; free; user-friendly for beginners).

3D model sharing and archiving: MorphoSource (Duke University; free for researchers; recommended for open deposition of 3D models with versioning and DOI assignment; supports OBJ, STL, PLY, and DICOM formats). Sketchfab (commercial; free tier; good for interactive 3D visualisation embedded in publications). Zenodo (CERN; free; not 3D-specific but suitable for metadata and analysis files accompanying MorphoSource deposits).