

Microtomographic investigation of a large corpus of cichlids

This manuscript ([permalink](#)) was automatically generated from habi/EAWAG-manuscript@dbbc7c0 on September 30, 2022.

Authors

- **David Haberthür**

 [0000-0003-3388-9187](#) ·  habi ·  [habi](#)

Institute of Anatomy, University of Bern, Switzerland

- **Mikki Law**

None

- **Kassandra Ford**

None

- **Marcel Hässler**

None

- **Ole Seehausen**

None

- **Ruslan Hlushchuk**

Institute of Anatomy, University of Bern, Switzerland

Abstract

A large corpus of Cichlids from Lake Victoria in Africa spanning a size range of 6 to 20 cm was nondestructively imaged using micro-computed tomography. The presented manuscript describes a method to efficiently obtain three-dimensional tomographic data sets of the oral and pharyngeal jaws and the whole skull of these fishes. We describe in detail how the data has been acquired to aid in reproducible research. The tomographic data we acquired (8.8 TB projection images) was reconstructed into 1.4 TB of three-dimensional images which are used for further projects. Herein we present our method and an outlook on two projects analyzing the acquired data; a morphological description of the oral and pharyngeal jaws of the fishes as well as a principal component analysis of landmark features on the fish skulls.

Introduction

History

- Cichlids from Lake Victoria
- Sample 'library' of EAWAG
- Valuable, hence non-destructive imaging is *paramount*

micro-CT

X-ray microtomography is a valuable tool to gain insights into the inner structure of very diverse samples, namely for specimens related to research done in the biomedical sciences. Namely in the 'fish sciences', X-ray microtomography has been employed as a method of choice to non-destructively assess the morphology of various samples [1]¹

Depending on the structures of interest biomedical samples are often tomographically scanned after the tissue/sample has been stained with a contrast agent, most often employing contrast agents containing heavy metals. Since the structures of interest for the two studies we touch upon in this manuscript (Cichlids teeth and skull bones) display large enough contrast to the surrounding tissue we did not stain our samples prior to the tomographic imaging presented here.

Materials and Methods

Sample procurement and preparation

The fishes were kept in 75% Ethanol for long-term storage in the EAWAG fish library. They were delivered to the Institute of Anatomy for X-ray microtomography investigation sorted into several batches by approximately equal length.

micro-CT imaging

All samples were scanned on two of the three available high-resolution 3D X-ray microtomography scanners of the Institute of Anatomy of the University of Bern in Switzerland, a SkyScan 1272 and a SkyScan 2214 (both Bruker microCT, Kontich, Belgium).

The fishes were sorted into ‘bins’ based on their physical size. We used a custom-made sample-holder to scan each of the fish in our machines. A sample holder was 3D-printed on a Form 2 desktop stereolithography printer (Formlabs, Somerville, Massachusetts, USA) and is freely available online [2] as part of a library of sample holders for tomographic scanning of biomedical samples [3]. The sample holder was custom-made for this project and is easily parametrized to the different width, height and length classes of the fishes.

In total, we acquired 340 tomographic scans of 127 different fishes. All the scanning parameters are collected in a table in the [Supplementary Materials](#), a generalized rundown is given below.

Since the fishes greatly varied in their length, the voxel sizes of each of the acquired datasets also varies greatly. We acquired datasets with (isometric) voxel sizes ranging from 3–50 µm.

Depending on the size of the specimen we set the x-ray source voltage to 50–80 kV and—depending on the voltage—to a current between 107 and 200 µA. Also depending on the size of the fishes, the x-ray spectrum was filtered either by an Aluminum filter of varying thickness (either 0.25, 0.5 or 1 mm) before digitization to projection images or recorded in an unfiltered way. In total we recorded 8.8 TB of projections images (*.?if files) for this project.

All the recorded projection images were subsequently reconstructed into a 3D stack of axial PNG images spanning the regions of interest of each fish. Since all the specimens were scanned with their mouths downward and rotating along their long axis, we manually rotated each of the reconstructed datasets so that the lateral axis through the fish was horizontal in relation to the x and y direction of each reconstructed slice. We reconstructed the projection images with NRecon (Version 1.7.4.6, Bruker microCT, Kontich Belgium) with varying ring artifact and beam hardening correction values, depending on each fish (again, all relevant values are listed in the [Supplementary Materials](#)). In total, this resulted in 1.4 TB of reconstruction images (*.rec*.png files).

While performing the work, a subset of the data was always present on the production system, for working with it (see [Preparation for analysis](#) below). A small bash script [4] was used to generate redundant (archival) copies of the raw projection images and copy all the files to a shared network drive on the `research_storage` infrastructure of the University of Bern, enabling easy collaboration on the data by all co-authors at the same time.

Data analysis

We wrote a set of *Jupyter* [5] notebooks with *Python* code to work with the images and wrangle the acquired data. The notebooks were written at the start of the project, to be able to process new scans as soon as they were reconstructed. Re-runs of the notebook added newly scanned and reconstructed fishes to the analysis, facilitating a an efficient quality check of the scans and batched processing of the data.

All Jupyter notebooks for this work are available online [6].

Preparation for analysis

The main Jupyter notebook for this manuscript dealt with reading all log and image files and preparing images for quality checking and the next steps. Briefly summarized the below process was implemented.

At first, *all* log files of *all* the scans were read into a list (*all* data of the notebook was saved into a *pandas* [7] dataframe). This already enabled us to extract the specimen name and scan, since we performed several scans for each specimen, e.g. a low resolution scan with large field of view for the whole head and one or several scans in high resolution focusing on the region of the oral and pharyngeal scans. From the log files we extracted the relevant values for double-checking the necessary parameters of each scan. All relevant values for each scan were also saved into the dataframe and saved out to the aforementioned table in the *Supplementary Materials* at the end of each run of the notebook.

After several ‘sanity checks’ of the data, we used *Dask* [8] to efficiently access the very large total amount of axial reconstruction PNG slices for this project (in the end a total of nearly a million single slices). On average, each of the tomographic datasets contains around 3000 slices, so the total amount of data is much too large to keep in memory. The use of the *Dask* library facilitated efficient (so-called lazy) access to the huge amount of data on disk.

At first, we extracted the central view of each of the three axial directions of the datasets, e.g. the ‘anteroposterior’, ‘lateral’ and ‘dorsoventral’ view and either saved those to disk or loaded them if already generated in prior runs of the notebook. The notebook then also generated the maximum intensity projection (MIP) for each of the anatomical planes and either saved them to disk or loaded them from prior runs.

At the end of the notebook we performed a final sanity check on the MIP images. In this check we examined the mapping of the gray values of the raw projection images to gray values in the reconstructions, e.g. checked that no overexposed pixels are present in the MIP images, which is an efficient way of checking this, since the MIP images have already been generated in prior steps of the notebook.

Image processing

Extraction of oral and pharyngeal jaws, visualization of tomographic data

To extract the oral jaw (OJ) and pharyngeal jaw (PJ) of the fishes, we used *3DSlicer* (Version 4.11.20210226) [9] extended with the ‘SlicerMorph’ tools [10] which help biologists to work with 3D specimen data. The reconstructed .png stacks were loaded into ‘ImageStacks’, depending on their size we reduced the image resolution (e.g. downsampled the images) for this first step. The three-dimensional volume was rendered via *VTK GPU Ray Casting*. A custom-made volume property (created by Kassandra Ford) was used as an input to view the scans. Using toggles in the volume rendering, we defined regions of interest (ROIs) for both the OJs and PJs in each specimen.. These

ROIs were then extracted in their native resolution from the original dataset for further processing. Using the grayvalue thresholding function in Slicers Segment Editor the teeth in both the oral and pharyngeal jaws were extracted. We used the ‘Scissor’ and ‘Island’ tools of the Segment Editor to isolate single regions.

Processed regions of interest were exported as NRRD [11] files. The three-dimensional visualizations of all regions of interest for each specimen were compiled into overview images (see Figure 1 for an example from the compilation document). In total we compiled overview of XXX specimens with full head morphology, oral jaw and lower pharyngeal jaw profiles.

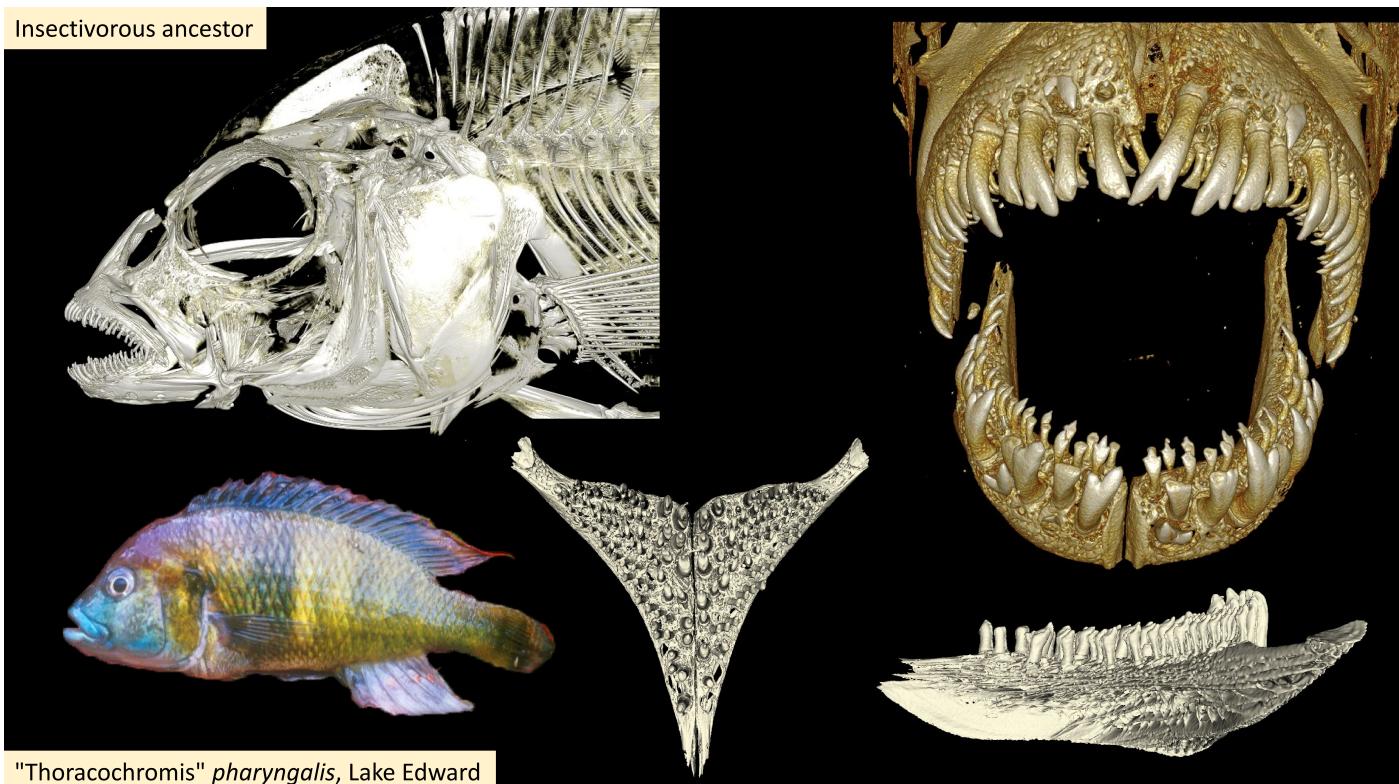


Figure 1: Overview of data from ‘KAT-13, Lake Edward, “Thoracochromis” pharyngalis (pharyngeal mollusc crusher – shrimp)’

Principal component analysis of skull landmarks

- Very superficial description of principal component analysis (?) from Kassandra. We do *not* want to cannibalize her upcoming manuscript, but only hint at what will be done.

Automatic extraction of otoliths

- Notebook: EAWAG/ExtractOtoliths.ipynb
- MIPs are oriented *anteroposterior*, *lateral* and *dorsoventral*
- Simple grayvalue plot along the longest axis
- Find peak of this grayvalue
- otolith is around maximum gray value along fish, (see Figure 2)

10619/head_rec_unbinned_17.5um

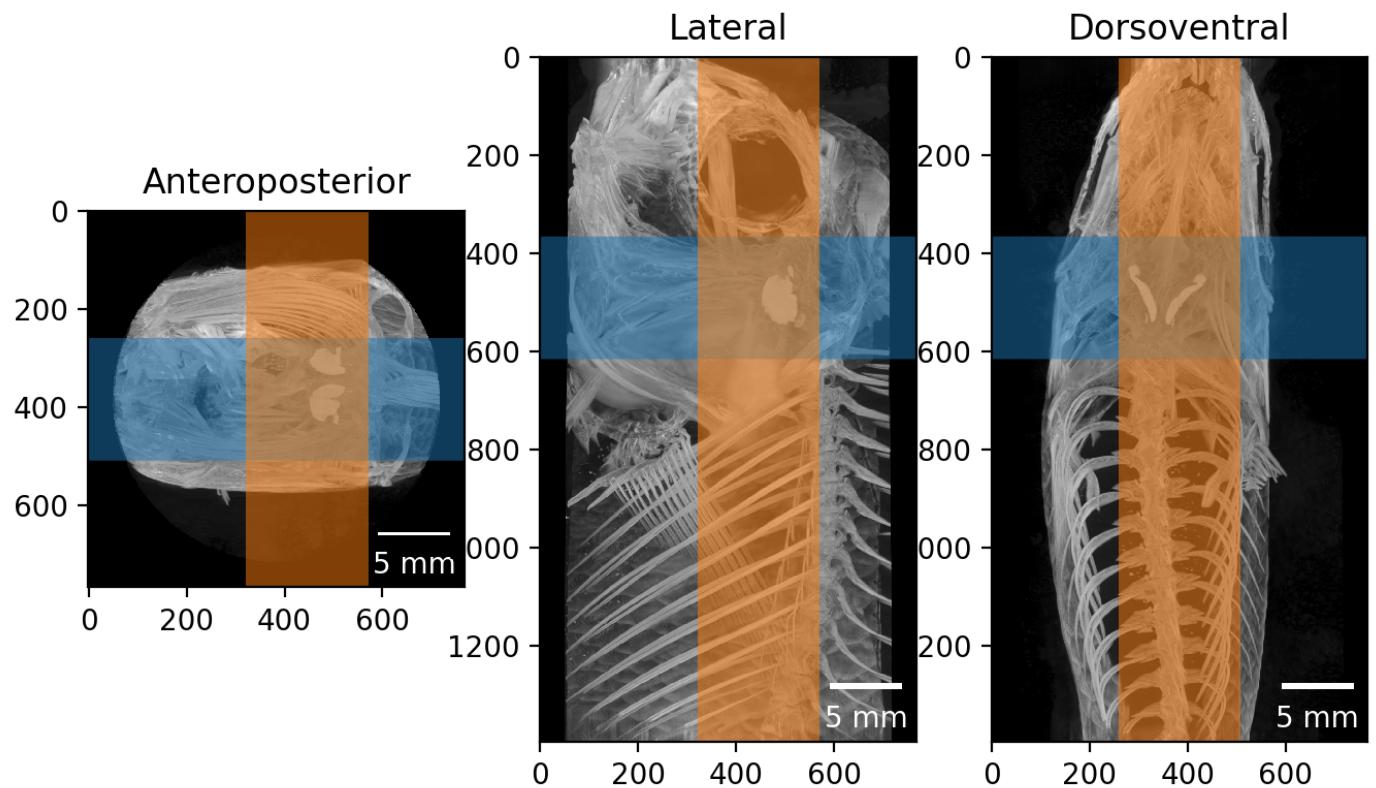


Figure 2: Automatic otolith extraction.

Results

- A lot of fishes
- A lot of scans
- A lot of data

Automatic extraction of otoliths

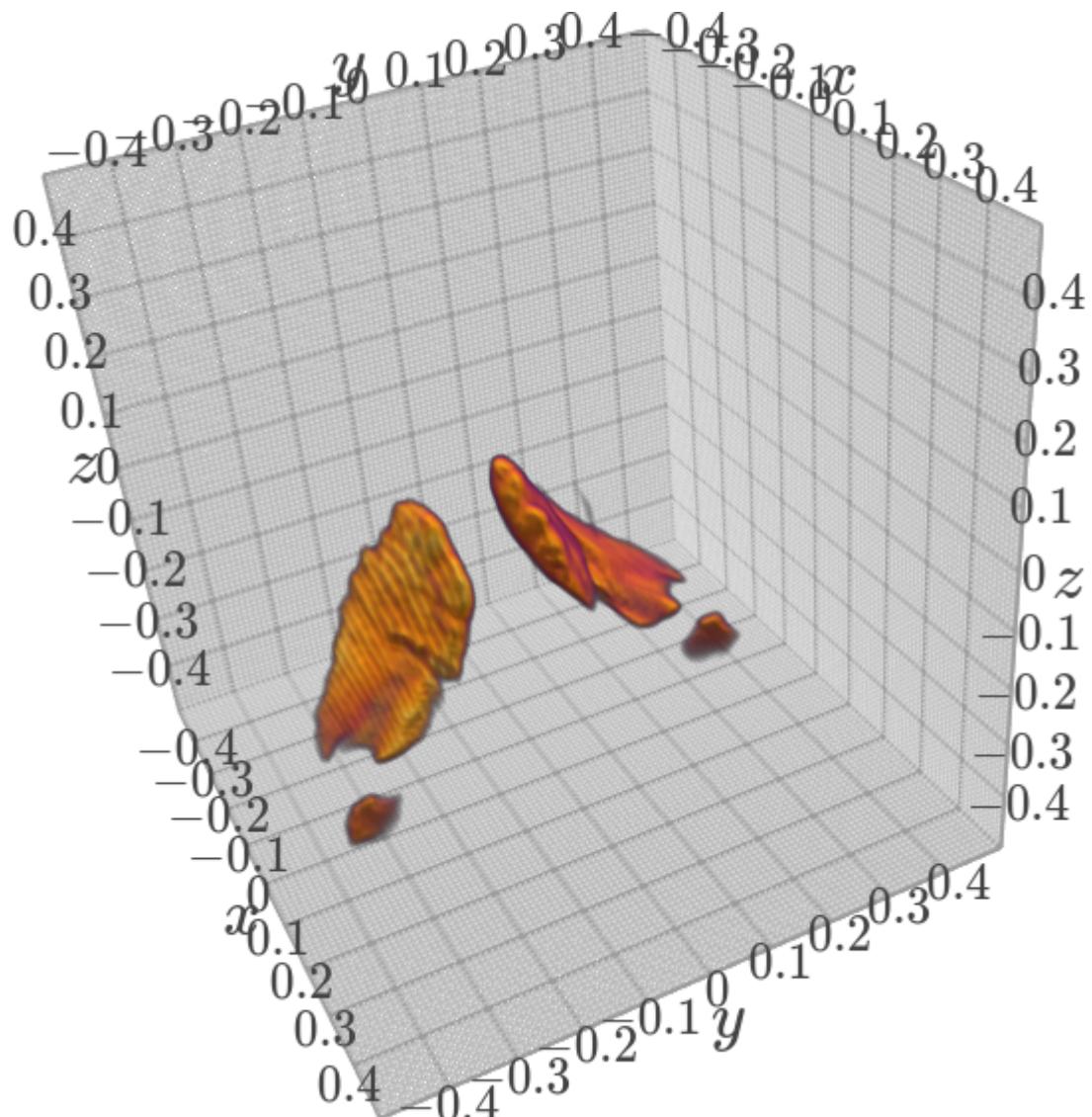


Figure 3: Result of automatic otolith extraction. Three-dimensional view of extracted otolith.

Discussion

The discussion of the results and the outlook to what we'll do in the future is going into this file here.

Acknowledgments

We thank the `manubot` project [12] for helping us write this manuscript collaboratively.

Supplementary Materials

Parameters of tomographic scans of all the fishes

The CSV file [ScanningDetails.csv](#) gives a tabular overview of all the (relevant) parameters of all the scans we performed. This file was generated with the [data processing notebook](#) and contains the data which is read from *all* the log files of *all* the scans we performed. A copy of the log files is available in a [folder in the data processing repository](#).

References

1. **CT Scans - #ScanAllFish**
Adam P Summers
(2015-09-29) <https://osf.io/ecmz4/>
2. **Hol3Drs/EAWAG.Fish.stl at master · TomoGraphics/Hol3Drs**
GitHub
<https://github.com/TomoGraphics/Hol3Drs>
3. **TomoGraphics/Hol3Drs: A release**
David Haberthür
Zenodo (2019-03-08) <https://doi.org/gg9fxh>
DOI: [10.5281/zenodo.2587555](https://doi.org/10.5281/zenodo.2587555)
4. **EAWAG/rsync-fishes.sh at master · habi/EAWAG**
GitHub
<https://github.com/habi/EAWAG>
5. **Jupyter Notebooks – a publishing format for reproducible computational workflows**
Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica Hamrick, Jason Grout, Sylvain Corlay, ... Jupyter development team
IOS Press (2016) <https://eprints.soton.ac.uk/403913/>
DOI: [10.3233/978-1-61499-649-1-87](https://doi.org/10.3233/978-1-61499-649-1-87)
6. **habi/EAWAG: First release, for minting a Zenodo DOI**
David Haberthür
Zenodo (2022-07-05) <https://doi.org/gqgdtp>
DOI: [10.5281/zenodo.6798632](https://doi.org/10.5281/zenodo.6798632)
7. **pandas-dev/pandas: Pandas**
The Pandas Development Team
Zenodo (2022-09-19) <https://doi.org/gqw4db>
DOI: [10.5281/zenodo.7093122](https://doi.org/10.5281/zenodo.7093122)
8. **Dask: Library for dynamic task scheduling**
Dask Development Team
(2016) <https://dask.org>
9. **3D Slicer as an image computing platform for the Quantitative Imaging Network**
Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona Fennessy, Milan Sonka, ... Ron Kikinis
Magnetic Resonance Imaging (2012-11) <https://doi.org/gfghgd>
DOI: [10.1016/j.mri.2012.05.001](https://doi.org/10.1016/j.mri.2012.05.001) · PMID: [22770690](#) · PMCID: [PMC3466397](#)
10. **SlicerMorph: An open and extensible platform to retrieve, visualize and analyse 3D morphology**
Sara Rolfe, Steve Pieper, Arthur Porto, Kelly Diamond, Julie Winchester, Shan Shan, Henry Kirveslahti, Doug Boyer, Adam Summers, AMurat Maga
Methods in Ecology and Evolution (2021-07-22) <https://doi.org/gqtgv8>
DOI: [10.1111/2041-210x.13669](https://doi.org/10.1111/2041-210x.13669)

11. <https://w.wiki/5mBK>

12. **Open collaborative writing with Manubot**

Daniel S Himmelstein, Vincent Rubinetti, David R Slochower, Dongbo Hu, Venkat S Malladi, Casey S Greene, Anthony Gitter

PLOS Computational Biology (2019-06-24) <https://doi.org/c7hp>

DOI: [10.1371/journal.pcbi.1007128](https://doi.org/10.1371/journal.pcbi.1007128) · PMID: [31233491](#) · PMCID: [PMC6611653](#)

1. For which David made a tomographic scan of an adult zebrafish ages ago. ↵