PaleoScan: Low-Cost Easy-to-Use High-Volume Fossil Scanning

Cláudio T. Silva
Yurii Piadyk
João Rulff
Daniele Panozzo
Maria Beatriz Silva
New York University

Antonio A. F. Saraiva
Naiara Cipriano Oliveira
Renan A. M. Bantim
Universidade Regional do Cariri

Flaviana Jorge de Lima
Universidade Federal de Pernambuco

Otávio J. F. Gomes
Universidade Federal do Ceara

Akinobu Watanabe
New York Institute of Technology
American Museum of Natural History

ABSTRACT

Fossils are crucial for understanding our natural history and the digitalization of fossils has paved the way for paleontologists to share and study them in greater detail. Yet, many fossil-dense regions, in particular low- and middle-income countries, lack the resources to digitalize their vast collections.

This project reports on a collaboration between paleontologists and computer scientists to design, build, and operate a device that can be deployed in the field for digitizing a collection of thousands of fossils. We introduce PaleoScan, a user-friendly, cost-effective, high-volume scanner designed to expedite the digitization of extensive fossil collections. PaleoScan is a self-contained 3D scanning system consisting of a light and compact mirrorless camera, a microcontroller, a ChArUco calibration board, and user-controlled LEDs. Software and data processing is cloud-based, where the user interacts with the system through a web application.

We deployed PaleoScan in a museum in Brazil with a world-class fossil collection. Our early results reveal its potential to revolutionize the scanning process for fossils.

CCS CONCEPTS

• **Applied computing**; • **Human-centered computing**;

KEYWORDS

digitalization, 3D scanning, 3D reconstruction, paleontology

1 INTRODUCTION

Paleontology is the study of ancient life generally relating to organisms that lived before or during the last Ice Age (ca. 12,000 years ago). Fossils, comprising 'body fossils' (remains of prehistoric organisms) and 'trace fossils' (evidence of life external to the organisms, such as footprints and fossilized dung), represent direct evidence of life forms that existed in the past. Our understanding of ancient biota is directly dependent on the quantity and quality of the fossil record. Fossilization is a rare event and a substantial portion of existing fossils are inaccessible or destroyed by erosion or metamorphism before they can be observed. Thus, the fossil record is incomplete, generally for organisms that lived earlier in Earth's history. As such, it is crucial that any fossils that are discovered are recorded, adequately characterized, and shared with the scientific community.

The digitization of paleontological collections can thus bring many benefits. Besides serving as a permanent record of the specimens, digitized collections can be accessed by anyone, anywhere in the world, making it possible for the community to collectively fill the gaps in our knowledge of biodiversity. These digital representations can be used to conduct research on a variety of topics, such as evolution, ecology, and climate change, as well as to educate the public about biodiversity and to inspire the next generation of scientists.

While fossils are found all over the world, they are common in low- and middle-income countries, which typically have limited resources to record, characterize, and share fossil data with the scientific community. As a result, these countries are excluded and unable to contribute to the global digitization effort.

Fossil Digitization For All. The digitization of fossils is a time-consuming and costly endeavor, which requires expensive equipment and trained personnel. To empower low- and middle-income countries to digitize their fossil collections, we need low-cost equipment that can be deployed in remote locations as well as equipment that is designed to be operated by personnel with limited computing expertise.

In October 2021, we started a collaborative project that brings together paleontologists in Brazil and computer scientists in the USA with the goal of developing technology that supports low-cost scanning for large fossil collections, such as the one in the Museu de Paleontologia Plácido Cidade Nuvens (MPPCN), which houses a world-class collection of fossils from the Araripe Basin in
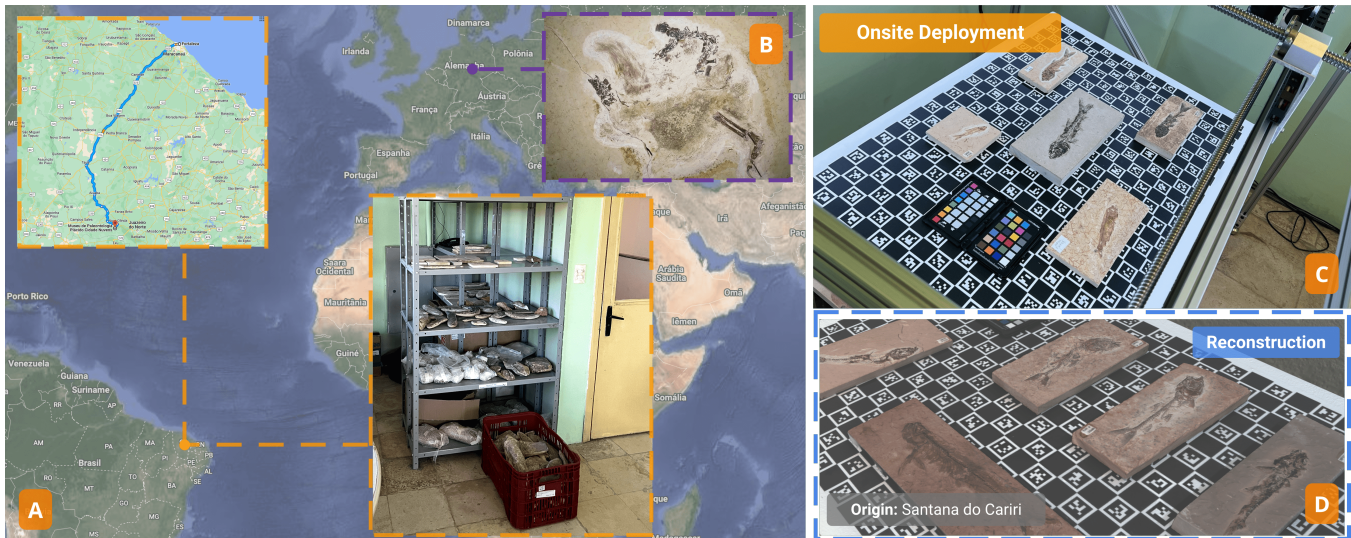
Figure 1: PaleoScan was designed to be a low-cost, easy-to-use, fossil digitization device able to scan hundreds of fossils a week, suitable for deployment in low-resource, rural zones, such as the one depicted above in Santana do Cariri, Brazil. The Museum of Paleontology Placido Cidade Nuvens houses a world-class fossil collection of enormous scientific importance. Our work makes it possible for this collection to be digitized in-situ, without the cost and risk of transporting the fragile fossils for scanning in other facilities. A-The location of the museum where the PaleoScan was deployed. B-*Ubirajara Jubatus,* one of the fossils recently reincorporated into the MPPCN collection after a long period in Germany. C-A batch of fossils loaded to PaleoScan. D-A reconstruction of one of the scans containing five *specimens.*

Northeastern Brazil. We started by gathering requirements for an ideal system, which should: 1) scan tens to hundreds of fossils of varying sizes a day; 2) generate scans that include high-resolution photography and depth information suitable to create 3D models of the surface of the fossil under controlled lighting; 3) be self-contained and not require a computer to be operated; 4) have an intuitive interface that does not require 3D scanning or imaging expertise; 5) operate without the need of a network connection. We refer to Section 4.1 for a detailed explanation of the importance of these requirements.

Over the next several months, we designed and built the hardware and software for the PaleoScan, and iteratively refined the system. In July 2023, we installed the first version of PaleoScan at MPPCN. The system has two main components: PaleoScanner, the scanner hardware, and PaleoDP, the software used to process, curate and annotate the scans, and derive the fossil images. The PaleoScanner, shown in Fig. 1 and described in detail in Section 4, is self-contained and requires only a connection to a power outlet to operate. It includes: a light and compact mirrorless high-resolution camera that can be automatically moved and positioned along a plane (two axes) and triggered to take pictures on demand; and a microcontroller with an attached touchscreen that can be used to operate the scanner. The specimens are placed on top of a ChArUco calibration board [13], which functions as a tool for automatic calibration and an implicit ruler. To control the lighting conditions of each scan, the system includes user-controlled LEDs in the perimeter. The scans are saved on an SD card, which the operator installs at the beginning of the day and removes at the end.

The processing of the raw data is handled by the PaleoDP tool (Section 5), independently from the acquisition, in a cloud infrastructure. After the data is uploaded from the SD card to the cloud, PaleoDP pre-processes the data and, through a web-based interface, allows users to annotate and curate the fossil images. At the end of the process, images are ready to be shared or uploaded into public databases.

Contributions. Our main contributions can be summarized as follows:

- We discuss the requirements for and present the design of PaleoScan, a low-cost and easy-to-use digitalization system for fossils. It was designed for flexibility, accuracy, and high throughput, with the goal of scanning large collections of fossils and other specimens in low-resource environments, including museums and excavation sites in low- and middle-income countries.
- We detail our implementation of the hardware and software components of PaleoScan.
- We describe the deployment of the PaleoScan in the Museu de Paleontologia Plácido Cidade Nuvens, a museum that houses a world-class collection of fossils in the Cariri region of Brazil.
- As part of two case studies, we demonstrate the effectiveness of our system, and example uses for paleontologists.

2 RELATED WORK

Researchers increasingly rely on digitized data of fossil and other biological specimens for conducting paleontological and evolutionary
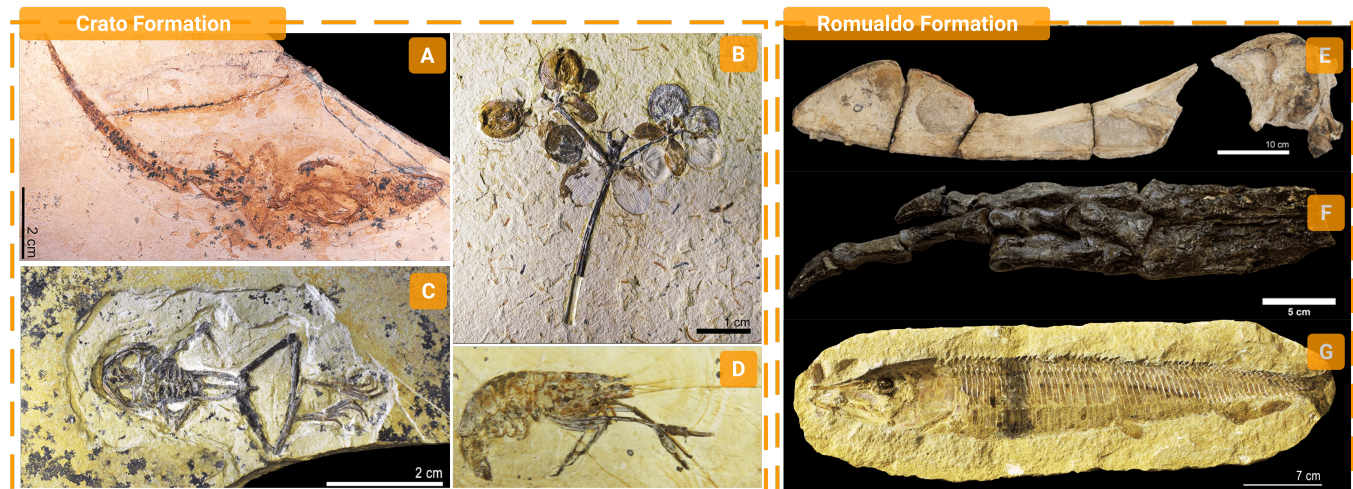
Figure 2: Examples of fossils from the collection of Museum of Paleontology Placido Cidade Nuvens. A-*Tijubina pontei*, collection number: MPSC-V010, Crato Formation. B-*Arariphrynus placidoi*, collection number: MPSCAm893, Crato Formation. C-*Cariria orbiculiconiformis*, collection number: MPSCPL3510, Crato Formation. D-*Beurlenia araripensis*, collection number: LPU1673, E-*Maaradactylus kellneri*, collection number: MPSCR2357, Romualdo Formation. F-*Aratasaurus museunacionalli*, collection number: MPSCR2089, Romualdo Formation. G-*Vinctifer comptoni*, collection number: MPSCP1015, Romualdo Formation.

studies. Unified databases of the digitized fossil record have become crucial tools in paleontology. Since the late 1990s, the Paleobiology Database (PBDB) [22][1] has remained an indispensable resource for fossil occurrence data, and providing tools for visualization and analysis related to animal diversity through time, geographic distribution, and other types of applications [32]. For those seeking fossil image data, databases such as MorphoSource contain 2-D photographs and 3-D models of fossils and other biological specimens organized by research projects. In recent years, MorphoSource has emerged as a pre-eminent data repository for 3-D tomographic data and their derivatives (e.g., 3-D mesh models) [4, 21][2]. Its continuously expanding database comprises phenotypic data of biological specimens, including fossils. Collectively, these online databases have substantially expanded and accelerated the accessibility of fossil image data across the scientific community.

While there are a number of online databases for storing and sharing phenotypic data of fossils, the resources required for high-resolution, 2-D and 3-D imaging of specimens has been a critical bottleneck. Generating high-quality 2-D images of fossils is tricky and time-consuming due to complexities of performing correct and consistent ligthing and focusing (see, e.g., [16]). The process is even more complex when fossils are very small or very large.

The standard modality for acquiring 3-D data of fossils and other whole-organism specimens has been (micro)computed tomography and surface scanners. The aforementioned devices are costly; scanning electron microscopy (SEM) and micro-CT scanners typically cost hundreds of thousands of U.S. dollars and portable laser surface scanners generally cost tens of thousands of U.S. dollars.

Even if newer and cheaper alternatives for 3-D scanning are becoming available, the data generated by these scanners often needs a substantial amount of manual processing using proprietary software, see, e.g. [10, 15, 33]. Furthermore, the rapid phenotyping of many specimens may require a dedicated facility and technical staff, which add to the overall cost of the imaging pipeline. Altogether, modern 3-D imaging capabilities are out of reach for many institutions with rich fossil collections, especially those in low- and middle-income countries.

3 DIGITIZING THE FOSSILS IN THE MPPCN

3.1 Background on the Cariri Region

The Cariri region, located in northeastern Brazil, is one of the world's foremost hubs for paleontological research. It hosts exceptionally well-preserved fossils from various areas of the Araripe Sedimentary Basin. This sedimentary basin comprises ten stratigraphic units that formed during the fragmentation of the Gondwana supercontinent and the subsequent opening of the South Atlantic Ocean [2]. The basin covers an area of about 12,000 km^2 and is geographically located in the southern part of the state of Ceará, the northwest of Pernambuco, and the eastern region of Piaui [8]. It is within this sedimentary basin that the famous fossils, preserved in laminated limestones of the Crato Formation and calcareous concretions of the Romualdo Formation, are found. Both formations are part of the Lower Cretaceous period and have significant geological importance [26].

The Crato Formation is known for its laminated limestones that contain many continental fossils, including arthropods, fish, amphibians, turtles, crocodylomorphs, pterosaurs, non-avian dinosaurs, and birds [26]. These fossils are commonly found and are

---

[1]https://paleobiodb.org
[2]https://www.morphosource.org

often well-preserved, exhibiting the soft tissues of the organisms. This formation also has an impressive diversity of mushroom and plant fossils, including ferns, gymnosperms, and angiosperms [8]. The fossils are frequently discovered as two-dimensional compressions or impressions [26], and their significance within paleontology is due to their abundance and excellent preservation [24, 26].

The Romualdo Formation encompasses a wide variety of rocks, including layered conglomerates, sandstones of different sizes, laminated limestones, marls, shales, coquinoid limestones, and coquinas [2]. The fossils of this formation are particularly noteworthy for their three-dimensional preservation within carbonate concretions [26]. Coprolites and fish fossils are the most commonly found fossils in these concretions [9]. However, traces of other creatures like pterosaurs, turtles, crocodylomorphs, dinosaurs, plants, gastropods, echinoderms, bivalves, and crustaceans can also be found within this formation [8, 26].

3.2 Digitization of the MPPCN Collection: Opportunities and Challenges

The MPPCN, localized in Santana do Cariri, Ceará, Brazil, houses a collection of about 11,000 fossils. Their collection consists of fossils primarily from the Crato, Ipubi, and Romualdo (Mesozoic Era), and Brejo Santo and Missão Velha (Paleozoic Era) formations [26]. Fig. 2 shows some examples of fossils from the Araripe Basin from the MPPCN's collection. The fossils in the MPPCN's collection were obtained from donations, private collections, seizure, rescue in mining areas, and controlled paleontological excavations–many of which were conducted by the Laboratory of Paleontology of Universidade Regional do Cariri (LPU) [26]. The MPPCN's collection is of immense scientific value as it offers an extensive look into the past with fossils dating back to the Jurassic and Cretaceous periods. The collection includes rocks, minerals, plants, invertebrates, and vertebrates, making it a comprehensive and fascinating exhibit for anyone interested in the Earth's history. Digitizing the collection at MPPCN epitomizes the challenges that accompany global digitalization. Despite its enormous scientific importance and invaluable collection, the museum is in a very low-resource, rural zone. The city of Santana do Cariri is estimated to have 17,726 inhabitants and is about a two-hour drive from the nearest urban centers, Crato and Juazeiro do Norte. Due to this distance and the large amount of fossils, it is risky to transport fragile fossils for scanning in other facilities. Thus, digitizing MPPCN's vast collection should be done on-site, where there is limited access to technicians with computational training, a reliable wireless network, powerful PCs, and other resources that are common in high-income countries.

4 PALEOSCAN HARDWARE

4.1 Design Criteria

We used the digitization challenges faced by the MPPCN to guide the design of the PaleoScanner. Given the size of their fossil collection, it is critical to have a *high-throughput* scanner. Since the museum is in a remote location without a reliable connection to the Internet or access to computers, *the scanner must be a self-contained device* and a low-cost mechanism is needed to *transfer the acquired data to an external computing infrastructure.* The MPPCN and other paleontological sites are harsh environments with high levels of dust and often hot weather. Consequently, the device must be *robust and withstand varying environmental conditions* and *easily repairable.* Last but not least, since the museum is resource-constrained, it is also crucial for the device to be low-cost.

There are also requirements vis-a-vis the data. An important requirement to catalog and share fossil data is the availability of metadata. The metadata is crucial for advanced analyses, such as the reconstruction of the 3D shape of the fossil. Therefore, we include the requirement for the software to provide an easy and efficient way to include information, such as fossil's formation, collection number, age, location, etc., by the expert paleontologists.

We devised further design specifications for the PaleoScanner based on the properties of our target objects, namely ancient fossils. Since fossils are preserved in rock form, their appearance is always somewhat textured. Moreover, fossils tend to have a distinct shape: since they are often found after a rock is cut, fossils usually have a flat face containing a sliced view of the specimen. Along with fossil shape, we considered the distinct classes of rock formations that are common for certain types of fossils. Most discovered fossils are small and stored in rectangular-shaped "rock frames." However, the sizes of individual fossils may vary significantly, with some exemplars reaching 1 meter in size. Additionally, since fossil sizes vary greatly, the device must accommodate both large and small specimens.

There are commercial 3D scanners available, such as GOM ATOS [31][3] or Artec Ray [30][4], that meet some of the individual design objectives; however, they often require manual actuation, a controlled lab environment, or are prohibitively expensive. We refer the reader to [35] for an excellent overview of structured light scanning methods. Assembling a good quality 3D scanner out of off-the-shelf parts has been a challenging but achievable task [25] with resources (including code) available for DIY solutions [17]. In our design, we capitalize heavily on the textured appearance of fossils and use a multiview stereo depth-like approach to reconstruction. This is different from, for example, structured light scanning since we are only using a single camera for both 2-D image and 3-D acquisition. There is no auxiliary projector used to "create" a textured appearance artificially. The fossils are naturally textured and we only need bright LED lights for low-noise image acquisition.

4.2 High-Throughput Scanning

Scanning a large number of fossils of varying sizes requires the data acquisition process to be as automated as possible. However, this process must also be non-intrusive and non-destructive since fossils are unique and rare artifacts. We, thus, implemented moving parts in the PaleoScanner to reduce the workload on the scanner operator. For one, we used a high-resolution optical camera operating in the visible light range for our data acquisition process. For the maximum level of detail, we employed shorter focal length lenses (17 and 11 mm), which provide a wider field of view. This wider field of view becomes a key facet in our optimization of the digitization process for many reasons. For example, it enables higher spatial resolution since lenses of shorter focal lengths can more easily focus at closer distances and a broader range of angles from which

---

[3]https://mfg.trimech.com/3d-scanner/zeiss-gom-atos-q/
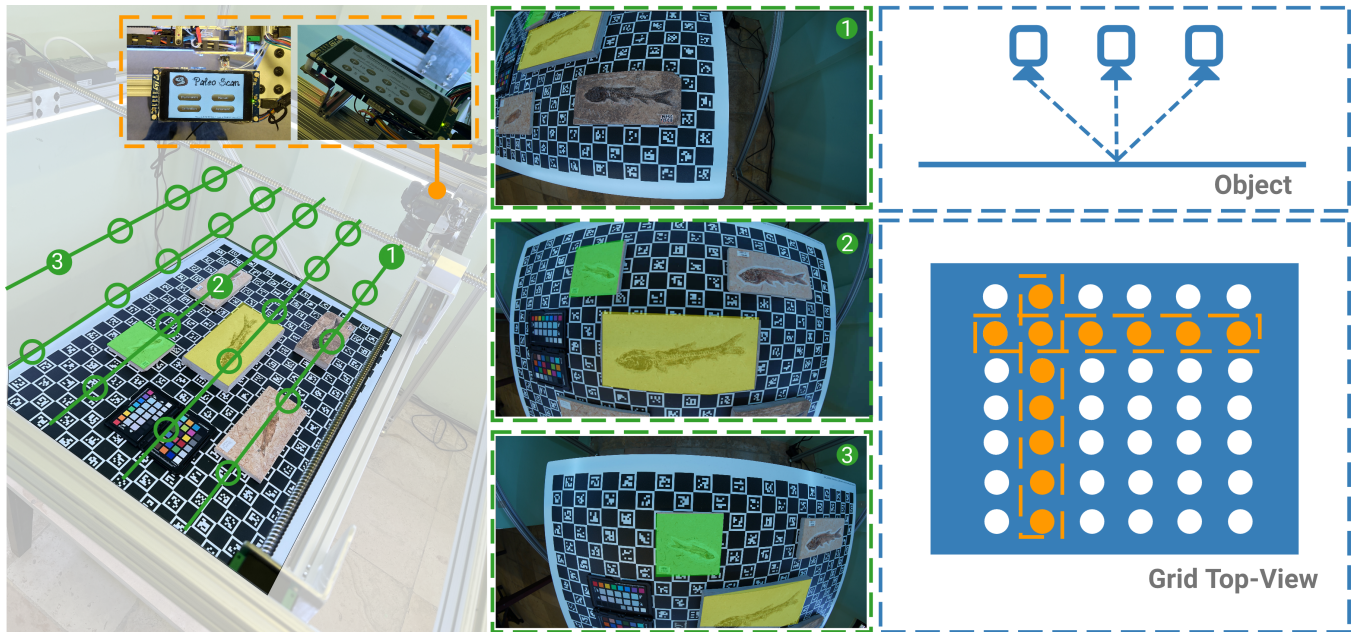[4]https://mfg.trimech.com/3d-scanner/artec-ray-2/

Figure 3: The PaleoScanner (shown on the left) works by taking pictures in a 2D pattern. The system was designed in such a way that the pictures overlap (middle). This way, each image taken can be used for multiview analysis of multiple fossils in its field of view (right). The fossils are also kept in focus by moving a vertically oriented camera parallel to the supporting table. The calibration board is an integral part of the system design enabling highly accurate computation of the camera location, thus enabling high-quality shape reconstruction.

the camera can view a given point on the fossil's surface as it moves (which aids 3D shape reconstruction).

It is important to point out that there is no need to scan all sides of each sample since there is typically only one side containing the fossil carveout. Therefore, the PaleoScan takes pictures from only one side of the specimen and relies on the camera's wide field of view to obtain sufficient coverage of other parts of the fossil from different view angles (Fig. 3, right). There is also no need to rotate the cameras in such a setup, which enables us to automate the camera movement using a linear stage while keeping the fossils in focus for the vertically oriented camera with a focus plane parallel to the supporting table.

For more advanced types of analysis, such as 3D shape reconstructions, it is essential to photograph the same part of the fossil from multiple view angles. Common approaches, such as structure-from-motion methods or structured light scanning, require dozens of images taken around each scanned object [11]; however, such techniques would be prohibitively expensive in terms of scanning time and storage when acquiring thousands of fossils.

Fossils also vary in size from several centimeters to up to a meter. Therefore, we use an approach that allows us to organize smaller fossils into batches to speed up the scanning process and remove the need for frequent reconfiguration of the scanning setup. We further leverage the wide field of view of our camera so that images taken to the left of some fossils in the batch can be reused as images to the right for the neighboring fossils (see Fig. 3).

Although it would be sufficient to use a single linear stage for left-to-right camera actuation, there are multiple benefits to implementing a 2D gantry for this purpose (see details in Section 4). First, it increases the utilization of the camera's sensor by imaging fossils in an orthogonal direction from top to bottom and from left to right. Second, images taken in any direction relative to others can now be selected for 3D shape reconstruction algorithms that are sensitive to corner cases, such as high-contrast lines/edges that can be ambiguous to reconstruct along one direction but trivial in another. Furthermore, it enables more types of analysis, including material reconstruction, when combined with controlled lighting.

4.3 Robust Image Registration

Any advanced analysis of the data requires accurate calibrations of the critical components of the scanner. The most central one is the camera, which is modeled by its intrinsic and extrinsic parameters. The intrinsic parameters are defined by the camera sensor (fixed by choice) and the lens used (17 and 11 mm in our case). There exist standard methods for estimation of the camera's intrinsic parameters, which rely on a calibration object of known dimensions [36]. We opted for the higher quality ChArUco calibration board [13], which includes a checker pattern alongside the AruCo markers sintered on the aluminum honeycomb substrate.

The advantages of the ChArUco calibration board are twofold: (1) it enables higher quality calibration of the camera's intrinsic parameters when using shorter focal length lenses with more distortions (i.e. 11 mm in our case), and (2) it is advantageous to use it as a
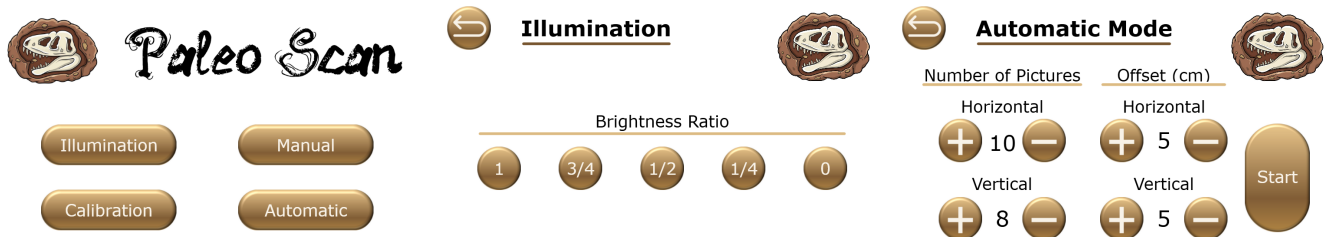
Figure 4: Selected screens of the PaleoScan GUI. Please note that the illumination brightness levels (middle) are achieved through spatial modulation (turning off every other LED) rather than temporal PWM to avoid flickering in camera images.

background when scanning the fossils since it enables refinements of the camera's extrinsic parameters, that is, it enables the accurate computation of the relative motion of the camera with respect to the board and fossils (see Fig. 8).

While it is possible to get approximate camera locations based on the linear stage dimensions, the accuracy is insufficient for 3D shape reconstruction tasks. This is due to the variability of camera positioning, which stems from the backlash and vibrations of the main frame when the Y-axis adjustment is active due to the increased moving mass that includes the X-axis linear stage. Temperature changes and possible collisions between the frame and the operator will also affect the geometry of the frame in unpredictable ways. We thus opt for refining the camera's extrinsic parameters using the ChArUco visual markers on each frame.

Another issue when using higher quality DSLR cameras equipped with a mechanical shutter is their inherent exposure time variability, which we estimated to be about 10% for our camera. That is, every time the picture is taken, the actual exposure might differ up to 10% from the set value because it is challenging for the manufacturers to achieve high reproducibility with a mechanical shutter. While not critical for regular photography, unevenly exposed neighboring images are challenging to use for material/shape reconstruction. To compensate, we use an X-Rite color checker placed alongside the scanned fossil for white balance and exposure/color calibration (Fig. 3, left).

4.4 Building the PaleoScanner

To build the PaleoScanner, one needs workshop/maker space facilities, which are typically available at US universities, including one soldering station, a bandsaw, a milling machine, a standard hand and power tools (i.e., a drill, screwdrivers set, etc.), multimeter, and a PC/laptop.

To the control logic and interface of the PaleoScan, we use a Discovery kit with STM32F469NI MCU [27][5]. It has a touch screen, which enables the development of an intuitive GUI (Graphical User Interface) to facilitate the control of all the components (Fig. 4). A detailed discussion of the GUI is described later in Section 4.5. It also includes pins and connections that can be used for reading and writing electrical signals, similar to an Arduino board. Our camera of choice is the SONY Alpha7 II. Its position on the frame is controlled by two stepper motors, which, in turn, are controlled by the microcontroller according to touch commands on the screen.

In addition to the motors, the microcontroller also controls the illumination and triggers the camera shutter. The Alpha7 II is a low-cost mirrorless camera. Such models tend to be smaller and lighter since they do not have the mirror box and the prism needed for a DSLR. For our particular application, the *live view mode* where we can see precisely the effect of changing settings in the resulting photograph greatly helps to make the system easier to use and setup. Other than these considerations, any camera with an audio input trigger function can be used in our system.

The microcontroller firmware was designed using TouchGFX Designer software [29][6] and STMCubeIDE [28][7] with a total development time of approximately 2 months. The main challenges in this design were ensuring that the stepper motors do not skip steps when moving the camera (along the Y axis in particular due to the added weight of the X axis linear stage), and adjusting the LED brightness without affecting the camera exposure (i.e. without flickering). The latter was achieved by turning every other LED on or off depending on the desired illumination brightness (a kind of spatial modulation) rather than using a common PWM (Pulse Width Modulation) approach.

4.5 Controller GUI

We designed the PaleoScanner's GUI to resemble a mobile device so that it would be simple to use without any technical training. The GUI has five screens that control the PaleoScanner: Home, Illumination, Calibration, Manual Mode, and Automatic Mode (see Fig. 4). The Home screen allows the user to navigate to the other menus. The illumination screen, which ideally is set first, has five options for the brightness level of the environment. After choosing the lighting, the user can return to the Home screen and, from there, go to the Calibration screen.

In the Calibration menu, the user can move the camera to the center of the frame and take pictures manually or automatically at regular time intervals. Calibration is a necessary step as it enables the computation of the camera's intrinsic parameters to be computed (see Section 5 for more detail). Every time the camera lens is changed or reconfigured (for instance, if the zoom level is changed) the PaleoScanner must be recalibrated. As we get more experience with PaleoScan, we have discovered alternative ways to obtain calibration information, including as a by-product of other operations; further details are provided later in the paper.

[5]https://www.st.com/en/evaluation-tools/32f469idiscovery.html

[6]https://www.st.com/en/development-tools/touchgfxdesigner.html
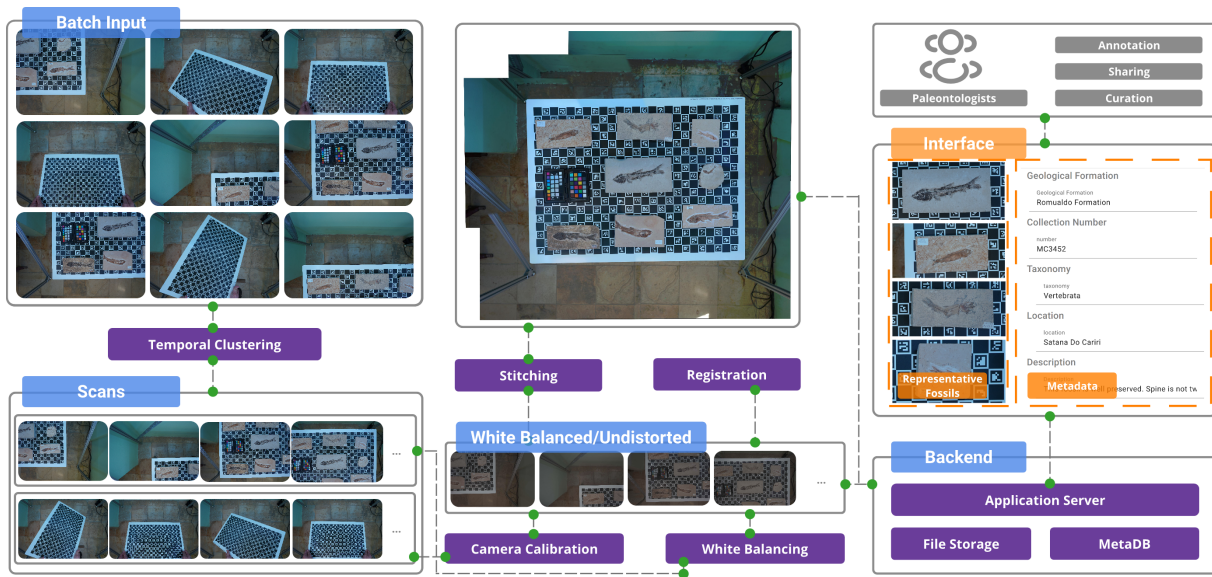[7]https://www.st.com/en/development-tools/stm32cubeide.html

Figure 5: The data processing pipeline implemented by PaleoDP is as follows. First, the uploaded batch of images undergoes an automated preprocessing step, including scan clustering, undistortion, and color correction. Secondly, the images are stitched, aiming to reduce the annotation effort of users to define fossils. Lastly, the images are saved and indexed in our backend infrastructure to support the PaleoScan interface, which Paleontologists will use to annotate relevant metadata.

After the setup, the user can start scanning the fossils. The Manual Mode screen allows pictures to be taken at user-specified positions and times, while the Automatic Mode screen requires an initial selection of a specific number of photos on a grid (photos by rows and columns) as well as horizontal (row) and vertical (column) offsets. It takes about 13 minutes to capture a grid of 10 by 5 photos (50 photos total) with a 5 cm grid size. The Automatic Mode is envisioned to be the primary way the system is used during production use, while the Manual Mode is convenient during the camera lens setup.

5 PALEOSCAN SOFTWARE: PALEODP

The primary goal of our project is to use imagery generated by the PaleoScanner to create a database containing photos of each individual fossil along with relevant metadata suitable for advanced analysis, including examples shown in Section 8. Our data processing pipeline developed to facilitate this consists of three major steps: (1) automatic preprocessing of the raw data captured by the scanner, (2) expert input by means of annotation of individual fossils, and (3) advanced processing that uses results the previous two steps as inputs. We describe each step in more detail below as also illustrated in Figure 5.

5.1 Preprocessing

Since the design of the PaleoScanner aims to maximize the throughput of scanned fossils, the raw images generated during each scanning round contain several fossils. Furthermore, multiple scans are performed in a row to minimize the overhead of extracting the data from the camera's SD card and uploading it to the cloud. Therefore, the first step in our processing pipeline is the segmentation of a

stream of images taken by the camera into individual scans based on the EXIF metadata contained by raw images, more specifically, the time when the image was taken. This process leverages the fact that image numbering is sequential and the time difference between pictures taken within the same scan round is much lower than the time between the last image of one scan and the first image of the following scan.

Once images have been grouped, the first dedicated set of calibration images is used to estimate the camera's intrinsic parameters. We use a well-established algorithm [36] to compute the camera's projection matrix K and distortion coefficients d values which required several dozen of the input images for good convergence. This step is unique in the sense that it can be performed using the JPEG images automatically generated by the camera software in addition to the RAW image data. For the actual scan, we are only using the RAW images processed identically after a complete set of calibrations was done.

Another important calibration that has to be performed is color correction. For that, we require a manual annotation of the corresponding white reference color square in the X-Rite color checker. Although manual, this step needs to be performed only once per lighting configuration, so the white balancing parameters W can be reused for multiple consecutive scans. An example is presented in one of our case studies (Section 8).

With all these calibration parameters known, we can then proceed with the preprocessing of raw images taken by the PaleoScan. First, the raw images are white-balanced, normalized for brightness, and gamma-mapped (we use gamma equal 2). Then the camera's intrinsic are used to undistort the images. Finally, we proceed to

Figure 6: The annotation interface of PaleoDP shows the representative crops of each defined fossil (left), all the automatically generated crops based on the bounding box created on the stitched image (middle), and the metadata associated with each fossil that the Paleontologists can edit (right).

the estimation of the camera's extrinsic parameters T (position) and R (orientation) for each image independently.

Estimation of the camera's extrinsic parameters is dependent on the use of the ChArUco calibration board as a background for fossils during scanning. It is important to leave at least two rows of markers visible between each fossil since two adjacent markers are used to estimate a single corner of the black-and-white checker. We are using SolvePnP function from the OpenCV library [5] for the estimation of T and R values for each image which completes our preprocessing step in the pipeline.

5.2 Annotation

User input is needed for the proper organization of the images with their corresponding metadata. Only paleontologists can provide such information as the fossil's age, the site where it was found, etc. We further capitalize on the opportunity of having an expert annotate the data to define the regions of interest (*ROIs*) around each fossil in at least one of the images for each scan. This information, with rough estimates of the fossil's height (H_{MIN} and H_{MAX}), will prove to be very useful for more advanced data analysis in the next step but we would like to first present the annotation tool that we developed (see Figure 6). Our tool consists of two main modules: a backend server and a user interface. The backend server contains a relational database (SQLite[23]), used to index the images and the corresponding fossils, and a file server where the actual images are stored. An application server handles requests through a REST API. The backend provides a user interface, built as a single-page application in Angular [18], with which the paleontologists interact. The interface is divided into two main views: the separation view and the annotation view. First, in the separation view, the goal is to allow the user to define regions where each fossil is contained. For this end, the following components are provided in the interface. On the left, is a list of all the images of a given scan (Figure 6-A), the user can select an appropriate image of the scan where the whole fossil of interest is visible. By clicking on it, a larger version appears on the annotation canvas (Figure 6-B) where the user can draw bounding boxes defining the region containing the fossil of interest. During this process, the user also sets a name for the fossil, which serves as an ID for the fossil in a relational database. Lastly, in the annotation

view, which has a similar design to the separation view, individual fossils cropped from the scans are available for the user to select and attach useful metadata to, such as where that fossil was acquired, estimated age, etc. Also, marking specific regions of the individual fossil is supported. The use of this feature is highlighted in Section 8.

5.3 Post-Processed Data and Applications

To recap, our original goal was to enable high-quality digitization of fossils as quickly and efficiently as possible. The pipeline described above starts with sequences of raw images from a large collection of fossils, and through a set of semi-automatic steps outputs high-quality imagery and associated metadata ready to be used in many downstream tasks. Here is the data that we have for each fossil: 1) we have *several* undistorted and color corrected *cropped* images from multiple positions; 2) for each of these images, we precisely know the camera location and parameters, allowing us to compute pixel correspondences between images; 3) we have manually annotated details entered by a paleontologist.

We can use this data for various applications. For instance, as we already mentioned and show later, we can do 3-D reconstruction from this data. Once we have enough annotated data, we should be able to train a machine learning model to automatically identify similar fossils instead of relying on the manual annotations. An alternative representation for the data would be to generate NeRFs (neural radiance fields) [20], and use that representation for generating novel views of the fossils. Once digitized, the possibilities are endless.

6 DEPLOYMENT OF THE PALEOSCANNER

6.1 Equipment Packing and Transportation

We originally envisioned constructing a small prototype to test in the field[8] before attempting to build a larger model, but it became clear that the benefits of a larger scanning area were too great to ignore. Not only did a larger scanner provide us with the ability to scan larger objects, but it also enabled us to scan multiple fossils at

---

[8]By *field* here, we refer to the place where fossils are prepared and handled such as university research labs or museums. Before proper preparation, fossils are not suitable for being digitized.

Figure 7: Deployment of the PaleoScanner required us to come up with a smart packing strategy that allowed us to safely carry the scanner to the field while enabling a streamlined reassembly process (See Section 6 for details). The PaleoScanner is currently deployed in the museum.

once, thereby significantly increasing our productivity. The main drawback of the larger scanner is its size and weight.

The issue of safely packing the PaleoScanner to survive a multi-stop plane trip followed by many hours of ground transportation was non-trivial. We decided to perform a partial disassembly, where the main operating scanning plane would be kept intact, but one scanning axis would be removed and transported separately in a tube enclosure. In order to protect the calibrated plane, we reused the legs to create a protective shell, which, when attached with metal pieces, made the whole setup rigid. A wood enclosure was custom-built to fit the frame. The sensitive electronic components, such as the camera and the microcontroller, were disconnected and packaged separately. The calibration board fit nicely in the middle of the protective shell.

After it was packed, we transported the PaleoScanner to the airport, only to be told that our package was too heavy to be shipped as luggage. At 109 lbs, it was 9 lbs over the limit for the airline, so we had to abort taking it and re-evaluate our packing methods. We decided to move the calibration board to a separate box and to use lighter wood for the side panels, which brought us to within the allotted weight. Since we now had an extra box, we included some spare parts in case something got damaged during shipping. After the multi-stop plane journey, we drove nearly 550 km (about 8 hours) to Santana do Cariri. In order to fit the packages in a small truck, we had to remove the outer wooden packaging.

6.2　Scanner Installation

The process of assembling the PaleoScanner was straightforward; most of the process required a single tool since we had used the same screw type nearly everywhere in the hardware. All in all, assembly took us roughly six hours. The main steps were as follows: we had to unpack all the parts, unscrew the legs which had been used as a protective shell for the main body, attach the legs and the axis that we removed for packing, and connect the camera and microcontroller. We had asked that a small wooden stump be built to support the calibration board, and this was prepared ahead of our arrival. This process is illustrated in Figure 7.

Towards the end of the assembly, there was a certain air of hesitation and anticipation of what would happen when we finally connected the PaleoScanner to power. Despite the extensive testing we conducted in our lab before transport, the potential for damage

while in transit was very real. To our genuine amazement, the system immediately came up. We selected a particularly beautiful fossil of a fish to be the first object that we scanned. By the time assembly was complete, it was late in the day, and other than focusing the lens, we did not perform any further calibration or optimization of the process.

7　SCANNER OPERATION & USER EXPERIENCE

While designing the PaleoScanner, we also co-designed a scanning workflow for the scanner operator to follow in order to capture all the essential measurements. Since scanning tens of thousands of fossils is very labor and time-intensive, even when the PaleoScanner can acquire dozens of smaller fossils at a time, ensuring the presence of all essential measurements is critical. The main steps of our proposed workflow are as follows.

We assume the PaleoScanner is deployed in proximity to the fossils. The overall process has two phases: the preparation phase, where the user sets up the PaleoScanner, and the acquisition phase, where the user scans the fossils. As part of the preparation phase, the user has to select the appropriately sized ChArUco calibration board to match the desired scanning height and fossil sizes accounting for the supporting table. In the system, we have two different ChArUco boards and the main difference between them is that one has a finer grid than the other, making it better-suited for smaller fossils. The next choice is to select the lens to be used such that they can focus at the desired scanning height and provide a wide field of view (that is, there should be overlap between images). Once the ChArUco board is placed on a table that will be used as a background, the camera has to be focussed to the appropriate height to enable objects to be placed on the board to be well focused. In order to have sufficient depth of field for the variety of fossil sizes that are about to be scanned, the user has to favor a smaller aperture on the lens. We note that the *live view mode* on the camera makes all these steps fairly intuitive and easy to perform. The final step in the preparation phase is to acquire a sequence of images of the calibration board at various angles. This is to be performed without the supporting table. This data will be used for the intrinsic parameters calibration.

After the process described above, the user should fix the illumination conditions at a desired level and put back the supporting
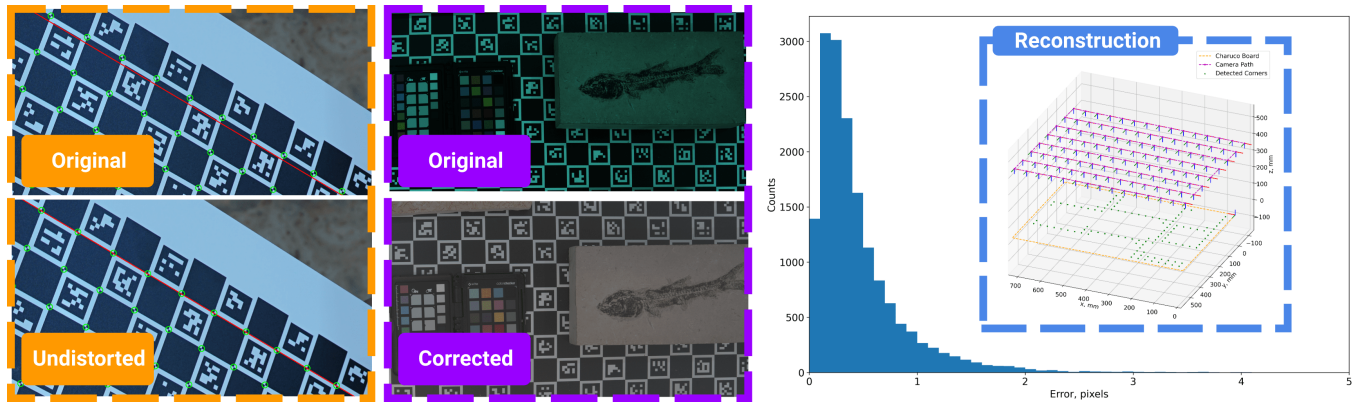
Figure 8: This figure shows some of the intermediate steps of the processing pipeline an image acquired with the PaleoScan goes through. On the left is the difference between a distorted image and an undistorted version. In the middle, we show the difference between the raw image and its color-corrected counterpart. Lastly, we highlight the reprojection errors of our intrinsic camera parameters estimation and reconstructed camera poses throughout the scan (right).

table, ChArUco board, and X-Rite color checker target for white balance calibration. At this point, the user can proceed with scanning of fossils. In fact, if the height and other settings are not changed, there is no *theoretical* need to repeat the preparation phase. In reality, it is good practice to redo the preparation phase at regular intervals to keep the overall system as calibrated as possible.

7.1 Usability Notes

Since the installation of PaleoScan at MPPCN, the system has been operated solely by the paleontologists and their students. In total, four people have used the system (3 female and one male). The training for the four users of PaleoScanner was straightforward. Our users consisted of two undergraduate students pursuing a Biological Sciences degree, a graduate student in Biological Diversity and Natural Resources, and a doctoral-level researcher in the field of Paleontology. All the users have no or limited computing experience. All of them were able to use the PaleoScanner repeatedly without any difficulty. The process for using the PaleoScanner was as follows: the researcher was responsible for camera configuration and capturing the images of the calibration board, and the students were responsible for setting up the GUI, selecting fossils, and scanning. All users reported that it was easy to use the PaleoScanner. When asked about the potential for the system, the researcher says that this type of equipment is fundamental for digitizing the MPPCN's collection, which will enable the dissemination of the Paleontology of the Araripe Basin, and promote the Cariri region.

At this time, PaleoScan has been used to digitize over 230 fossils of varying sizes. The images are uploaded into the cloud and processed as outlined in Sec. 5. We have not changed the software that runs on the actual hardware. One issue that we have encountered is related to the acquisition of the calibration images. In some of the scans, the calibration images are taken too close to the camera, and the board is not completely visible from the pictures. We have addressed this problem by recording a demonstration video that shows the correct way to take the calibration images. We have dubbed the video in Portuguese and made it available to the team

at MPPCN. As we will see later when we discuss the 3-D reconstruction, we have discovered that we can perform good calibration directly from the scans, potentially making the calibration images redundant. At this time, we still believe the calibration images are useful, and given that it only takes about 5 minutes to acquire them, we are keeping them on our protocol.

Another issue that came up was the arrangement of the fossils in the ChArUco board. Sometimes, two fossils can be too closely packed, making it harder for us to separate them automatically. We have included instructions on this issue too.

At this time, we rely on the operators to check the quality of the images, and whether they adhere to our requirements. In the ideal world, as images are taken, they would be processed immediately and feedback would be given to the operator (for instance, the system could say: "calibration board too close to the camera"). In order for us to do this, we would need to process images on the fly and in situ, which is something that we do not currently do.

8 CASE STUDIES

In this section, we present two examples of advanced analysis performed on the data collected during the deployment of PaleoScan at the museum. They have collected over 1 TB of image data collected so far. We were able to scan over 230 fossils, small and large, and gained valuable insights into fossil's properties for the purpose of 3D shape reconstruction as well as texture-based identification of the fossil's origin.

8.1 Fossil reconstruction

PaleoScan was designed to generate data that can be used for 3-D reconstruction through a technique called Structure from Motion (SfM) [12]. In principle, SfM is a rather straightforward technique: given a set of images with known camera parameters and relative positions, we can intersect rays between corresponding points on the images to generate 3-D points. If we find a large set of said correspondences, we can compute a dense point cloud, which
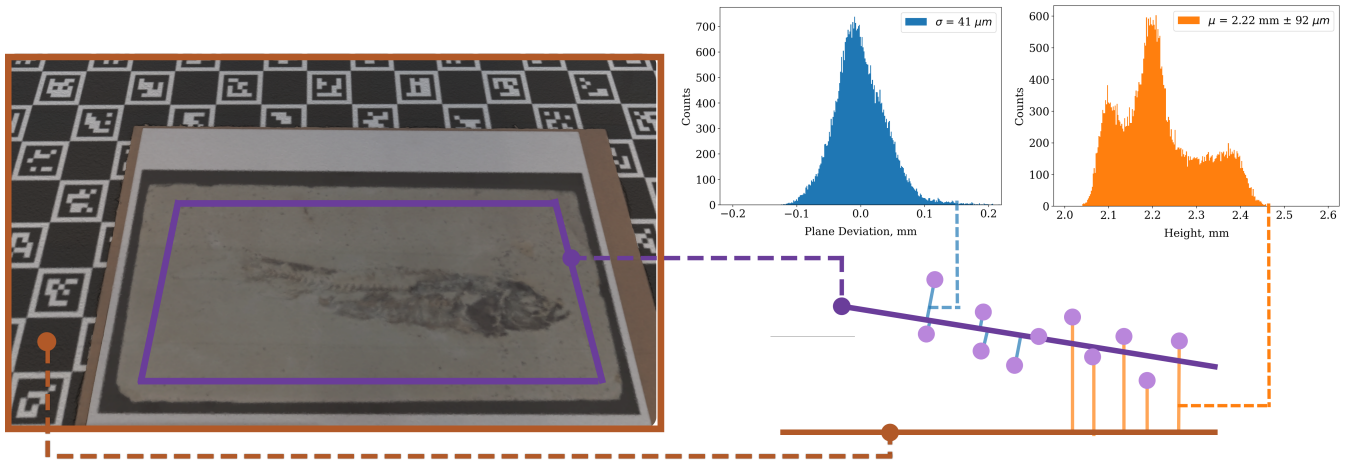
Figure 9: An artificial (fake) flat fossil (left) was prepared and scanned to test the accuracy of different reconstruction methods. The distribution of deviations from the fitted plane for the vertices of the fish region in the image, as reconstructed using Agisoft, is shown in the middle ($\sigma = 41\ \mu m$). The corresponding heights of these points above the recovered calibration board level are shown on the right (the average height = 2.22 *mm* $\pm$ 92 *μm*).

can be triangulated into a 3-D model [3]. There are excellent commercial and open-source implementations of SfM, and we have experimented with several of them, including code that we wrote ourselves. Two successful commercial products include 3DF Zephyr [34] and Agisoft Metashape [1].

Let us use our experience with 3DF Zephyr to demonstrate how the data can affect the performance of SfM. First, after taking several images of a fossil on a white table under natural lighting by holding the camera in our hand and allowing the camera to focus automatically, we attempted to perform reconstruction with 3DF Zephyr. Under these conditions, 3DF Zephyr was not able to determine camera parameters (which changed as the camera kept focussing from the different positions). The PaleoScan images, which are taken on a grid under controlled lighting on top of the ChArUco board, make the process much easier, and under these conditions, 3DF Zephyr has no problem finding the poses of images. 3DF Zephyr can optionally use external camera calibration information; in fact, we compared the quality of using the calibrations computed by 3DF Zephyr versus the PaleoScan specific calibrations (that we compute), and the quality of the reconstructions is higher with our camera calibration, although for many situations the default ones are still of acceptable quality. Our color-corrected images dramatically simplify finding correspondences.

To summarize, the undistorted and color-corrected images (Fig. 8) produced by our data processing pipeline, together with accurate camera information offered by the PaleoScan, make it much easier for SfM techniques to reconstruct fossil features on a submillimeter scale. We have been exploring a multiview variant of patch matching-based stereo depth reconstruction algorithm, which is very accurate, but it is a bit too conservative, and it misses many potential matches. Until we can make this technique work accurately and reliably across all situations, we have decided in our current data processing pipeline to use Agisoft Metashape [1]. Accuracy details are discussed a bit later in this section. We used a

15 X 8 scanning grid for most of our scans, offering 120 views of the scanned surface. The typical runtime when reconstructing fossils is between 15 to 45 minutes per set of images (that is, a collection of fossils) on a MacBook Pro M3 Max with 128GB of RAM and does not require any input from users.

Comparison to other approaches for 3-D reconstruction. The 21st century has witnessed incredible advances in digitizing biological specimens, including fossils. Figgins [10] article on "graphics in computational paleontology" provides a primer to the topic with excellent examples. Published during the early period of 3-D fossil digitization, Lyons et al [19] shows how to use a 3-D color laser scanner to create a highly accurate $100\mu m$-resolution model of the mosasaur *Tylosaurus*. The article mentions that the theoretical maximum resolution for the scanner is $10\mu m$ based on the manufacturer's notes. In their work, they report performing 30 scans of the fossil under different orientations, which were post-processed into a single model. They write: "Scanning the basisphenoid-basioccipital at $100\mu m$ required four hours, with an additional 3 hours of post-production time used to compile, assemble, and reduce the model." Their fossil is 11.3*cm* in length, smaller than most of the ones that we have scanned. In their paper, they demonstrate the usefulness of the 3D models by highlighting the ability of paleontologists to create a virtual copy of the fossil to share with colleagues around the world and manipulate the fossil in a 3-D digital environment to observe anatomical features. Johnson et al [15] is a more recent example using a hand-held laser scanner. They also performed the scans in multiple phases, each of which they estimated took 30 minutes. PaleoScan was designed to scan fossils of a certain type. This specialization enables us to streamline the whole operation compared to other existing options for digitizing biological specimens (e.g., Lyons et al [19] and Johnson et al [15]). We believe PaleoScan is easier to use as the users do not require any computing expertise. It is also efficient in terms of the user time; on average, it takes 16 minutes to take a full set of 120 images which can be used to digitize
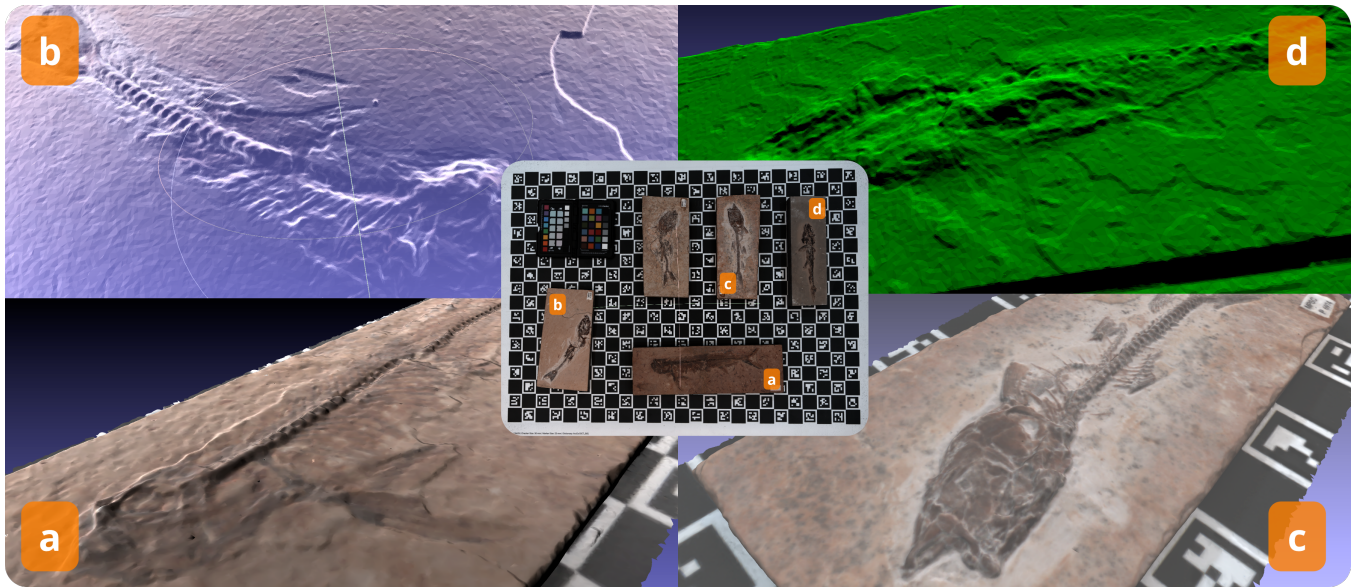
Figure 10: A reconstruction of an entire scan is shown in the middle. The corresponding fossils were rendered using different visualization methods to highlight the geometry and texture of the fossils. The clear bone structure of the fish spine shows good preservation of the original fossil.

many fossils at the same time (depending on fossil size), and we have automated almost all the data processing tasks. While prior work (e.g., Lyons et al. [19]) shows scans of single fossils, we have demonstrated the ability to scan hundreds of fossils in a relatively short amount of time.

Reconstruction accuracy. When deciding on a scanning and reconstruction method there is always a tradeoff in terms of scanning time/data size and reconstruction accuracy that needs to be made. In order to explore the different reconstruction methods, we manufactured a **fake** flat fossil by printing a high-resolution image of a fossil and gluing it on a flat substrate to serve as a benchmark with known ground truth geometry (see Fig. 9, left). We then scanned it by capturing 30 images in a 5 X 6 grid and applying our processing pipeline, including color correction and image undistortion. We used several 3D reconstruction codes on this data with the goal of finding the best technique and estimating its accuracy. At this time, Agisoft Metashape [1] produces the most accurate results (Fig. 9, middle). Of note is the residual tilt of the fake fossil when placed on top of the calibration board, which causes the asymmetric distribution of the heights for the reconstructed points (Fig. 9, right). Still, the standard deviation for the height estimation (i.e. the substrate's thickness) does not exceed 100 μm.

The importance of scanning fossils for paleontologists. The reconstruction of 3D shapes of fossils has become increasingly popular and critical in paleontology. By examining 3D images of fossils, researchers could gain biological insights based on the anatomical traits of fossil specimens. The collection of specimens provides an understanding of taphonomic processes and environmental conditions that have influenced the preservation of organic remains over time. This approach aids in identifying decomposition patterns, abrasion, fracturing, and other post-mortem processes, providing a

more comprehensive picture of the biology and ecology of ancient lifeforms.

The use of 3D images facilitates detailed comparisons among various fossil specimens and species, which are essential for identifying individual variations and differences between populations and even for discerning between related species. For instance, when analyzing 3D images obtained by scanning the fish species *Dastilbe crandalli*, variations in the preservation of individuals become evident, with the majority retaining articulated three-dimensional vertebrae. This suggests that the preservational environment was a low-energy setting with little disturbance, which played a role in maintaining the organism's natural joints. Nevertheless, some specimens exhibited displacement of the skull bones, with the skull moving to the dorsal region of the body, indicating body fragmentation before fossilization, which could be due to biotic or abiotic factors. All these taphonomic and preservational processes could be readily inferred by analyzing the three-dimensional images of many specimens from the same site generated by the PaleoScanner.

Fossil scanning is an essential tool in paleontology because this method also enables the creation of three-dimensional digital models, preserving the morphology and structure of specimens as faithful copies. It also fosters remote accessibility to fossils, thereby facilitating global collaboration among researchers. Digital reconstruction provides a detailed analysis that enhances our understanding of taxonomy, evolution, and the preservation of fossils. These reconstructions also serve as powerful tools for scientific dissemination and education, making paleontological discoveries more accessible to the public. Finally, fossil scanning offers a method of digitally preserving rare or fragile specimens, minimizing the need for frequent handling, and reducing the risk of damage. In summary, this technology represents a significant advance in the way
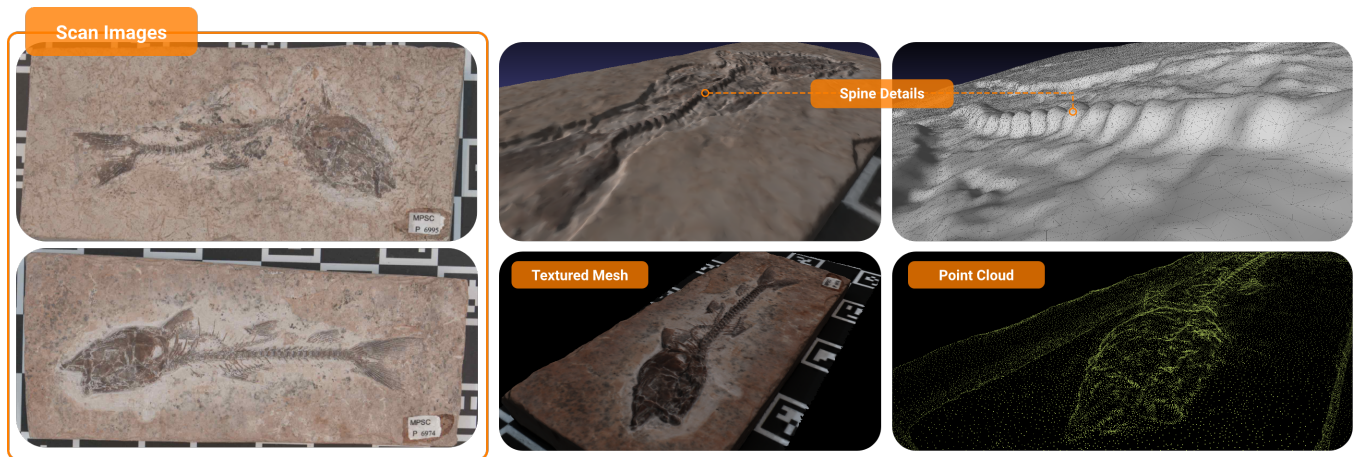
Figure 11: Sample reconstructions of the 3D shape of the fish fossils. On the left, are crops of fish scans used in the reconstruction process. The right-top shows a closer look at a well-preserved fish spine, suggesting a low-energy depositional environment, which helps maintain the organism's natural joints. Finally, different renders of the 3D shapes of the fish are shown on the bottom.

paleontologists conduct their research, promoting a more detailed, accessible, and preservative approach to the study of the remnants of the past.

8.2 Fossil trafficking

Recently, despite efforts to combat fossil trafficking in Brazil, this practice persists as an ongoing problem in the Cariri region. It is important to note that, in Brazil, it is illegal for individuals not affiliated with a national or state research institution to carry out fossil excavations without first obtaining authorization from Brazil's National Mining Agency [7, 14]. In addition, the Brazilian constitution states that fossils found in the country are state property, making it illegal to sell or export them without permission [6] [9].

The fossil richness of the Araripe Basin has attracted researchers from around the globe due to the remarkable diversity, quantity, and excellent preservation of fossils from various groups of animals and plants. Their deposits contain fossils from the end of the Jurassic to the Early Cretaceous period. Because of their scientific value, fossils have been illegally traded on the black market, sometimes for thousands of dollars. Fossils illegally smuggled out of the country cannot be used in scientific publications, as demonstrated by a recently retracted article in the Cretaceous Research journal about a theropod dinosaur that was illegally taken out of Brazil. (The specimen has recently been returned to MPPCN.) In Fig. 12, we show some examples of fake fossils donated to the MPPCN's collection and one cast fossil available for purchase by visitors at the museum.

Researchers from MPPCN and other Brazilian paleontologists have actively supported the combat against fossil trafficking in the Cariri region. Part of this effort includes providing training courses for law enforcement agencies to help them identify fossiliferous materials. Thanks to their efforts, hundreds of fossils extracted from the

Araripe Basin has been prevented from being trafficked. Therefore, the database produced by PaleoScan, installed at MPPCN, can be a helpful tool for law enforcement. By utilizing paleontological data from
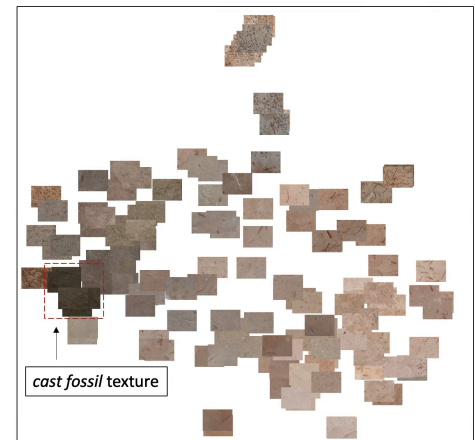
the Araripe Basin obtained with PaleoScan, agents will have easy access to images of real fossils, making it easier to differentiate them from casts and even fake fossils. Also, there are many opportunities for creating automatic approaches to aid law enforcement. For instance, on the right, we show a UMAP projection of patches of fossil textures. We highlight the texture from the cast fossil in the lower bottom corner. Even with the limited data that we have processed so far, we can speculate that creating databases of materials and textures is a promising approach to help the identification of fossils, and whether they are intact or have been tampered with.

9 DISCUSSION AND LESSONS-LEARNED

Plan evolution through requirement gathering and field visits. When we started the project and before visiting Santana do Cariri, we had a totally different plan for digitization. It involved much more infrastructure, and included purchasing expensive CT scanners, which are relatively very large, and also have much more

---

[9]http://www.planalto.gov.br/ccivil_03/decreto-lei/1937-1946/Del4146.htm

Figure 12: Cases of intentional adulteration of fossils from the Romualdo Formation by "peixeiros," aimed at making them appear unique and rare, have damaged the integrity of the material and obscured significant morphological characteristics. A-Indeterminate fish fossil transformed by hand into a lizard. B-Visitors to the MPPCN can purchase acrylic resin casts made with limestone powder of *Tharrhias araripis fish*. C-Fossil fish of the species *Vinctifer comptoni* partially complete, with the caudal fin carved into the rock; below the skull of the larger fish, a fragment of an undetermined fossil has been carved into the shape of a chelonian.

complex maintenance. Also, operators have to go through more complex training. After visiting the museum, and the excavation sites in October 2021, we started to brainstorm about other potential solutions that were more suitable for field deployment and also were "easier to use and operate" from a very broad perspective. We pivot to building an optical scanner, which as we gather requirements in multiple conversations and iterations converged to the design that we present in this paper.

Low-cost digitization through remote computation. As shown in this paper, PaleoScan is a low-cost scanning device that can scan fossils at a high rate. One of the reasons for the low cost, and also the simplicity of operation is that we push the most complex and costly computations to the cloud. This is particularly true for all the data processing tasks involved in shape reconstruction. This turns out to also have an advantage in terms of software organization, as we can improve capabilities remotely.

What would we do differently?. During and after installation, we identified areas for improvements on the hardware. Some of these are just cosmetics, such as better hiding internal wiring from prying eyes who might be tempted to touch internal components. Although the equipment lives in a limited-access room in the museum, there is still quite a bit of human traffic there, and we believe it would be better to cover components up and to have a more "production"

look to the equipment going forward. Also, about half of the cost is the camera and lens. We believe these expensive components need to be easily detached to allow for safe storage when not in use in a locked cabinet (or similar).

Materials acquisition. Another improvement to PaleoScan is the acquisition of rock material properties, such as a bidirectional reflectance distribution function (BRDF). We used individually programmable LEDs (WS2812B[10]) to achieve spatially modulated brightness control that eliminates flickering in camera images. These LEDs could also be programmed to illuminate the scanned area from multiple sides to capture how light is being scattered by the fossils at different angles. The tradeoff is that this would increase the number of images that need to be taken per single camera position.

Other regions and facilities around the world. There are many regions around the world where PaleoScan can be deployed. For instance, The Research and Advancement Corporation for Atacama's Paleontology and Natural History (CIAHN Atacama) in Chile, appears to be suited for using a similar setup to the one that is currently deployed in Santana do Cariri.

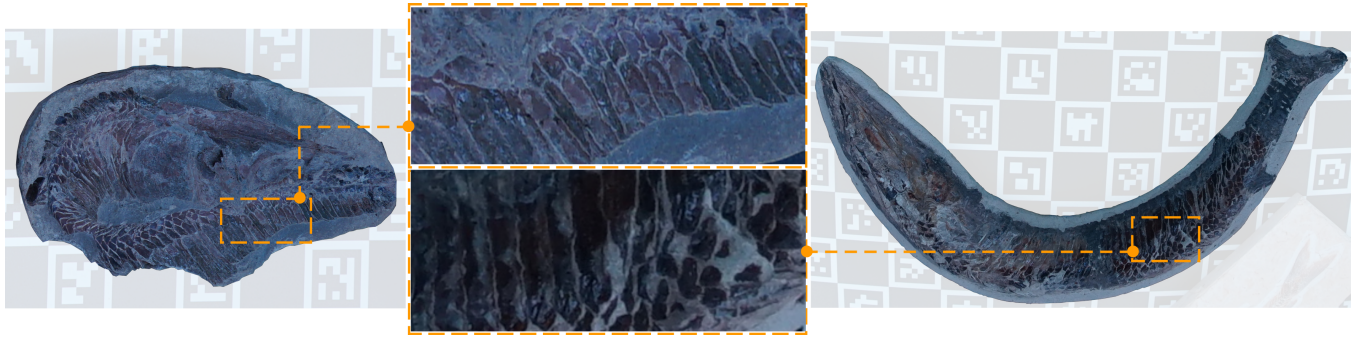[10]https://cdn-shop.adafruit.com/datasheets/WS2812B.pdf

Figure 13: This figure shows how the details of the acquired images can inform paleontologists regarding the preservation of the fossils by analyzing the morphology of their longitudinal scales.

Short-term goals. At this time, we are planning a digitalization sprint to scan 1,000 fossils at MPPCN. This will help optimize our protocols, and make sure the tools, interfaces, and data products generated best serve the scientific needs of paleontology.

10 FUTURE PLANS

PaleoScan provides a powerful, yet cost-effective, path towards rapid digitization of fossil and other specimens. These qualities unlock modern imaging capabilities in virtually any institution around the world, especially for those with fewer resources. This goal is important, not only for making imaging tools more accessible but also for ensuring that our collective knowledge of the fossil record is as complete as possible, without being hindered by inequalities and inequities in resources. Here, we have focused on the application of PaleoScan for fossil specimens, but the technology could also aid in rapid digitization efforts of other objects, including archaeological artifacts and artwork.

The images produced by PaleoScan can help paleontologists infer useful information regarding the preservation of the fossils. Images like the *Vinctifer comptoni*, shown in Fig. 13, allow detailed observation of the morphology of their longitudinal scales, which are coated with a layer of ganoin. This further underscores the remarkable quality of preservation found in the Romualdo Formation. The availability of this type of data in a public database allows any researcher who has access to the images to utilize their high-resolution capacity, which in turn enables accurate reproduction of the body morphology of the organisms. This detailing is crucial for research in the taxonomy and systematics fields.

11 ACKNOWLEDGMENTS

REFERENCES

[1] Agisoft. 2023. Agisoft Metashape User Manual. Last accessed on 12/12/2023. https://www.agisoft.com/pdf/metashape-pro_2_0_en.pdf
[2] Mario Luís Assine, JA de J Perinotto, M Andriolli Custódio, VH Neumann, FG Varejão, PC Mescolotti, et al. 2014. Sequências deposicionais do andar Alagoas da Bacia do Araripe, nordeste do Brasil. *Boletim de Geociências da PETROBRAS* 22, 1 (2014), 3–28.
[3] F. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, and G. Taubin. 1999. The ball-pivoting algorithm for surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics* 5, 4 (1999), 349–359. https://doi.org/10.1109/2945.817351
[4] Doug M Boyer, Gregg F Gunnell, Seth Kaufman, and Timothy M McGeary. 2016. Morphosource: archiving and sharing 3-D digital specimen data. *The Paleontological Society Papers* 22 (2016), 157–181. https://doi.org/10.1017/scs.2017.13
[5] Gary Bradski. 2000. The openCV library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer* 25, 11 (2000), 120–123.
[6] Brasil. 2023. DECRETO-LEI Nº 4.146, DE 4 DE MARÇO DE 1942. http://www.planalto.gov.br/ccivil_03/decreto-lei/1937-1946/Del4146.htm.
[7] Juan C Cisneros, Aline M Ghilardi, Nussaïbah B Raja, and Paul P Stewens. 2022. The moral and legal imperative to return illegally exported fossils. *Nature Ecology & Evolution* 6, 1 (2022), 2–3. https://doi.org/10.1038/s41559-021-01588-9
[8] FJ De Lima, AAF Saraiva, and Juliana Manso Sayão. 2012. Revisão da paleoflora das formações Missão Velha, Crato e Romualdo, Bacia do Araripe Nordeste do Brasil. *Estud. Geológicos* 22 (2012), 99–115. https://doi.org/10.18190/1980-8208/estudosgeologicos.v22n1p99-115
[9] Emmanuel Fara, Antônio ÁF Saraiva, Diógenes de Almeida Campos, Joao KR Moreira, Daniele de Carvalho Siebra, and Alexander WA Kellner. 2005. Controlled excavations in the Romualdo Member of the Santana Formation (Early Cretaceous, Araripe Basin, northeastern Brazil): stratigraphic, palaeoenvironmental and palaeoecological implications. *Palaeogeography, Palaeoclimatology, Palaeoecology* 218, 1-2 (2005), 145–160. https://doi.org/10.1016/j.palaeo.2004.12.012
[10] Stephen Figgins. 2001. Graphics in computational paleontology. *IEEE Computer Graphics and Applications* 21, 6 (2001), 6–13. https://doi.org/10.1109/38.963454
[11] Jan-Michael Frahm, Pierre Fite-Georgel, David Gallup, Tim Johnson, Rahul Raguram, Changchang Wu, Yi-Hung Jen, Enrique Dunn, Brian Clipp, Svetlana Lazebnik, et al. 2010. Building rome on a cloudless day. In *Computer Vision–ECCV 2010*. Springer Berlin Heidelberg, Berlin, Heidelberg, 368–381. https://doi.org/10.1007/978-3-642-15561-1_27
[12] Yasutaka Furukawa, Carlos Hernández, et al. 2015. Multi-view stereo: A tutorial. *Foundations and Trends® in Computer Graphics and Vision* 9, 1-2 (2015), 1–148. https://doi.org/10.1561/0600000052
[13] Sergio Garrido-Jurado, Rafael Muñoz-Salinas, Francisco José Madrid-Cuevas, and Manuel Jesús Marín-Jiménez. 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* 47, 6 (2014), 2280 – 2292. https://doi.org/10.1016/j.patcog.2014.01.005
[14] Elizabeth Gibney et al. 2014. Brazil clamps down on illegal fossil trade. *Nature* 507, 20 (2014), 20–20. https://doi.org/10.1038/507020a
[15] R. A. Johnston, K. Barnes, T. Lovell-Smith, and N. Brent Price. 2004. Use of a Hand-held Laser Scanner in Palaeontology : A 3-D Model of a Plesiosaur Fossil. Akaroa, New Zealand. https://polhemus.com/_assets/img/Use_of_a_Hand-held_Laser_Scanner_in_Palaeontology.pdf
[16] Jack Kallmeyer. 2018. Photographing Your Fossils. Last accessed on 12/8/2023. http://drydredgers.org/photographing_your_fossils.htm
[17] Douglas Lanman and Gabriel Taubin. 2009. Build your own 3D scanner: optical triangulation for beginners. In *International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH ASIA 2009, Yokohama, Japan, December 16-19, 2009, Courses Proceedings*. ACM, Yokohama, Article 2, 94 pages. https://doi.org/10.1145/1665817.1665819
[18] Mattias Levlin. 2020. DOM benchmark comparison of the front-end JavaScript frameworks React, Angular, Vue, and Svelte. Available at https://urn.fi/URN:NBN:fi-fe2020051838212.

[19] Patrick D Lyons, Marc Rioux, and R Timothy Patterson. 2000. Application of a three-dimensional color laser scanner to paleontology: an interactive model of a juvenile Tylosaurus sp. basisphenoid-basioccipital. *Palaeontologia Electronica* 3, 2 (2000), 16.

[20] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 2021. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. *Commun. ACM* 65, 1 (dec 2021), 99–106. https://doi.org/10.1145/3503250

[21] MorphoSource. 2023. MorphoSource. Last accessed on 9/13/2023. https://www.morphosource.org/

[22] NSF. 2023. The Paleobiology Database. Last accessed on 9/13/2023. https://paleobiodb.org/#/

[23] Mike Owens and Grant Allen. 2010. *SQLite.* Apress LP New York, New York.

[24] Alita Maria Neves Ribeiro, Yong Yang, Antônio Álamo Feitosa Saraiva, Renan Alfredo Machado Bantim, João Tavares Calixto Junior, and Flaviana Jorge de Lima. 2023. Arlenea delicata gen. et sp. nov., a new ephedroid plant from the Early Cretaceous Crato Formation, Araripe Basin, Northeast Brazil. *Plant Diversity* 87 (2023), 149–200. https://doi.org/10.1016/j.pld.2023.06.008

[25] Claudio Rocchini, Paolo Cignoni, Claudio Montani, Paolo Pingi, and Roberto Scopigno. 2001. A low cost optical 3D scanner. *Comput. Graph. Forum* 20, 3 (2001).

[26] AAF Saraiva, FJ de Lima, OA Barros, and RAM Bantim. 2021. Guia de fósseis da Bacia do Araripe. *Governo do Estado do Ceará, Ceará-Brazil* 0, 0 (2021), 376.

[27] STMicroelectronics. 2022. STM32F469NI. Last accessed on 9/14/2023. https://www.st.com/en/evaluation-tools/32f469idiscovery.html#overview

[28] STMicroelectronics. 2022. STMCubeIDE. Last accessed on 9/14/2023. https://www.st.com/en/development-tools/stm32cubeide.html

[29] STMicroelectronics. 2022. TouchGFX. Last accessed on 9/14/2023. https://www.st.com/en/development-tools/touchgfxdesigner.html

[30] Trimech. 2023. Artec Ray. Last accessed on 9/14/2023. https://trimech.com/products/artec-ray-3d-scanner

[31] Trimech. 2023. GOP ATOS. Last accessed on 9/14/2023. https://trimech.com/products/gom-atos-q-12m

[32] Mark Uhen, Bethany Allen, Noushin Behboudi, Matthew Clapham, Emma Dunne, Austin Hendy, Patricia Holroyd, Melanie Hopkins, Philip Mannion, Phil Novack-Gottshall, et al. 2023. Paleobiology Database User Guide Version 1.0. *PaleoBios* 40, 11 (2023). https://doi.org/10.5070/P9401160531

[33] Ray Wilhite. 2003. Digitizing large fossil skeletal elements for three-dimensional applications. *Palaeontologia Electronica* 5, 2 (2003), 4.

[34] 3DF Zephyr. 2023. 3DF Zephyr, The Complete Photogrammetry Solution. Last accessed on 12/12/2023. https://www.3dflow.net/

[35] Song Zhang. 2018. High-speed 3D shape measurement with structured light methods: A review. *Optics and Lasers in Engineering* 106 (2018), 119–131. https://doi.org/10.1016/j.optlaseng.2018.02.017

[36] Zhengyou Zhang. 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 11 (2000), 1330–1334. https://doi.org/10.1109/34.888718