Applications

Nathalie Klug*, Markus Kramer, Fahri Mazrek, Lorenz Wührl, Hossein Shirali, Rudolf Meier and Christian Pylatiuk

Automated photogrammetric close-range imaging of small invertebrates using acoustic levitation

Automatisierte photogrammetrische Bildgebung kleiner Wirbelloser unter Verwendung akustischer Levitation

https://doi.org/10.1515/auto-2025-0139
Received December 22, 2025; accepted April 20, 2026

Abstract: Photogrammetry is gaining popularity in biodiversity research as it provides multiple views of the same specimen. However, handling the small specimens often requires fixing them with pins to position them in front of a camera, which can damage them. Acoustic levitation offers a solution by enabling lightweight objects to be suspended in air without physical contact. We demonstrate how small insects can be imaged from different perspectives using acoustic levitation to position and rotate them in mid-air. As opposed to other levitation systems with rotation control, prior knowledge of the geometry of the levitated object is not required. This way, automated multi-view imaging of insects with various morphologies is possible. To show the images are suitable to create a 3D representation, an image set is reconstructed in the 3D space using neural radiance fields (NeRF).

Keywords: acoustic levitation; biodiversity; digitization; multi-perspective imaging; neural radiance fields; photogrammetry

Zusammenfassung: Photogrammetrie gewinnt in der Biodiversitätsforschung zunehmend an Bedeutung, da sie mehrere Ansichten desselben Exemplars ermöglicht. Die Handhabung kleiner Exemplare erfordert jedoch oft, dass sie mit Nadeln fixiert werden, um sie vor einer Kamera zu positionieren, was zu Beschädigungen führen kann. Die akustische Levitation bietet hier eine Lösung, indem sie es ermöglicht, leichte Objekte ohne physischen Kontakt in der Luft schweben zu lassen. Wir zeigen, wie kleine Insekten aus verschiedenen Perspektiven abgebildet werden können, indem wir sie mithilfe der akustischen Levitation in der Luft positionieren und drehen. Im Gegensatz zu anderen Levitationssystemen mit Rotationssteuerung sind keine Vorkenntnisse über die Geometrie des schwebenden Objekts erforderlich. Auf diese Weise ist eine automatisierte multiperspektivische Bildgebung von Insekten mit unterschiedlichen Morphologien möglich. Um zu zeigen, dass die Bilder für die Erstellung einer 3D-Rekonstruktion geeignet sind, wird ein Bilddatensatz mithilfe von Neural Radiance Fields (NeRF) im 3D-Raum rekonstruiert.

Schlagwörter: akustische Levitation; Biodiversität; Digitalisierung; multiperspektivische Bildgebung; neural radiance fields; Photogrammetrie

***Corresponding author: Nathalie Klug**, Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344, Baden-Württemberg, Germany, E-mail: nathalie.klug@kit.edu.
https://orcid.org/0009-0007-8185-8779
Markus Kramer, **Fahri Mazrek**, **Lorenz Wührl**, **Hossein Shirali and Christian Pylatiuk**, Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344, Baden-Württemberg, Germany
Rudolf Meier, Center for Integrative Biodiversity Discovery, Leibniz Institute for Evolution and Biodiversity Science, Museum für Naturkunde, Invalidenstraße 43, 10115, Berlin, Germany; and Institute for Biology, Humboldt University Berlin, Unter den Linden 6, 10117, Berlin, Germany

1 Introduction

Ultra close-range photogrammetry is used for the analysis of small objects in various fields, including forensics [1], cultural heritage [2], and biodiversity research [3]. To obtain high-quality 3D reconstructions and to make accurate analyses, photogrammetry requires high-quality images from multiple perspectives. This technology has especially gained popularity in biodiversity research since it is low-cost due to

advancements in imaging technology and computing power improvements [3]–[5].

In biodiversity research, insects have become a priority because they play a vital role, contributing more than 75 % of the species-level biodiversity and being responsible for many ecosystem functions [6]. One major obstacle to studying insects is the lack of user-friendly identification tools. Fortunately, various machine learning tools promise relief, as demonstrated in recent studies [7], but training the models will require sorting specimens to species with DNA barcodes [8] and then imaging them to have input data for training the models. Efficient tools for digitizing insects will thus be important and cost-effective methods would have the additional advantage of enhancing collaboration between biologists identifying new species, engineers developing high-throughput techniques, obtaining data for biodiversity of the tropics [9], and making new discoveries accessible to the general public.

Multi-perspective imaging with extended depth of field (EDOF) would be a major step in this direction because the external morphology of new insect species could be well illustrated, and closely related species could be more easily distinguished when images from multiple perspectives are available [3], [10], [11]. Furthermore, identification of the insects' gender and measurements of body parts, such as tibia length, could be obtained semi-automatically and be more accurate when multiple perspectives are available since parallax errors are avoided [3]. Collection of such trait data for species is increasingly important for holistic biomonitoring of biodiversity [12].

The digitization of insect samples with 2D images is feasible with various systems. The Entomoscope, an open-source photomicroscope, is a low-cost device for high-resolution imaging of insects [13]. The digitization is even becoming semiautomatic with robots such as the DiversityScanner, which sorts, images, and classifies insects preserved in ethanol [14]. However, all images show the insects from only one perspective. In contrast, the BIODISCOVER robot uses two cameras to take images of insects from multiple perspectives as they float down a cuvette [15]. However, the image quality is limited for taxonomic work, and EDOF imaging is not feasible. Finally, the DiversityScanner-360° allows automatic imaging from multiple perspectives perpendicular to the longitudinal axis of rotation [16] with EDOF of insects that are preserved in ethanol. However, more perspectives with EDOF are needed to get a complete picture of the insects, and dried specimens also have to be imaged.

Until now, systems that allow imaging from multiple perspectives around more than one axis of rotation require that the insects are dried and individually fixed with an insect pin before imaging [3]–[5], [11]. This step is time- and labor-intensive and cannot be easily automated. Furthermore, it is highly desirable to keep the insects in wet storage in ethanol for DNA preservation instead of drying and pinning them. The pin also covers parts of the insect during the imaging and damages the physical specimen. Additionally, all existing systems focus on larger insects (>5 mm) which can fit on a pin, although all dominant taxa of flying insects mostly consist of species that are much smaller [17].

Acoustic levitation uses sound waves to suspend small, low-density objects in air [18]. It has already found several applications, for instance, in the fields of pharmacy for the development and analysis of pharmaceuticals, e.g., the preparation of amorphous drugs [19]. This physical effect can therefore also be used to obtain high-resolution images of small invertebrates for photogrammetry and taxonomic analysis without the use of pins. Open-source acoustic levitation systems exist, which allow levitation, as well as position and orientation locking, of sub-wavelength objects of unknown geometry [20], [21]. Furthermore, systems with multiple phased array transducers, allow rotation control of levitated objects by controlling the transducers individually or in small groups [22], [23]. However, these systems require prior knowledge of the geometry.

Here, we introduce a system enabling automated photogrammetric imaging for the multi-view digitization of small insects with EDOF. It combines an automated acoustic levitation device with a camera system, overcoming the challenge of covering the entire object of interest with the camera's field of view without the interference of other objects, such as a pin. The invertebrates are trapped between high-pressure regions resulting from the acoustic waves, and rotated by generating and controlling the needed signals for the ultrasonic transducers with a Field Programmable Gate Array (FPGA). In contrast to previous levitation systems, we achieve rotational control of small objects without prior knowledge of their geometry. Macro photography is employed for high-resolution, multi-perspective, EDOF imaging. To showcase a potential use case the stacked images are combined into a 3D representation using neural radiance fields (NeRF).

The work is structured as follows. Section 1 motivates the topic and provides an overview of the current state of the art. Section 2 introduces the fundamentals of acoustic levitation and imaging. Section 3 presents the system design, imaging workflow, and experimental results, including representative images and a demonstration of 3D reconstruction. Section 4 summarizes the work and discusses

implications, limitations, and potential applications of the proposed approach.

2 Materials and methods

2.1 Acoustic levitation

Acoustic levitation is a method to suspend and move small, low-density objects in the air. For acoustic levitation to work, there must be an equilibrium of forces between the sound radiation pressure and gravitational force. The acoustic forces (F) are calculated using the negative gradient of the Gor'kov potential (U)

$$F = -\nabla U \tag{1}$$

$$U = 2K_1(|p|^2) - 2K_2\left(|p_x|^2 + |p_y|^2 + |p_z|^2\right) \tag{2}$$

$$K_1 = \frac{1}{4}V_P\left(\frac{1}{c_o^2\rho} - \frac{1}{c_P^2\rho_P}\right) \tag{3}$$

$$K_2 = \frac{3}{4}V_P\left(\frac{\rho_o - \rho_P}{\omega^2\rho_o(\rho_0 + 2\rho_P}\right) \tag{4}$$

where p is the complex pressure and p_x, p_y, p_z are the respective pressure gradients in the x, y, z directions. V_P is the volume of the particle, c_o is the speed of sound of the medium and c_P the speed of sound of the solid state particle. ρ_o is the density of the medium and ρ_P is the density of the solid state particle [20], [24].

Ultrasonic phased array transmitters with a frequency of 40 kHz are typically used, as the frequency is directly proportional to the sound radiation force, and thus higher forces can be generated as compared to 20 kHz transmitters [25]. The use of multiple transducers generates standing waves in which small objects can be suspended.

When controlling the ultrasonic transducers individually, different trap types can be implemented. Twin traps, also called acoustic tweezers, trap objects in a mirror-symmetrical pressure field [20], [23]. A simulation of such a field, created with the open-source Ultraino software [20], is depicted in Figure 1. Transducers of the same color in the image emit the ultrasonic waves in phase, while transducers in another color emit the ultrasonic waves with a phase shift.

When using systems with 40 kHz transmitters, the acoustic levitation is typically limited to objects of approx. 4 mm in size, as the object size then fits well with the dimension of the standing wave [20]. However, twin traps have the advantage of being able to levitate objects in the Mie Regime, meaning objects larger than the wavelength, with position and orientation control [23].

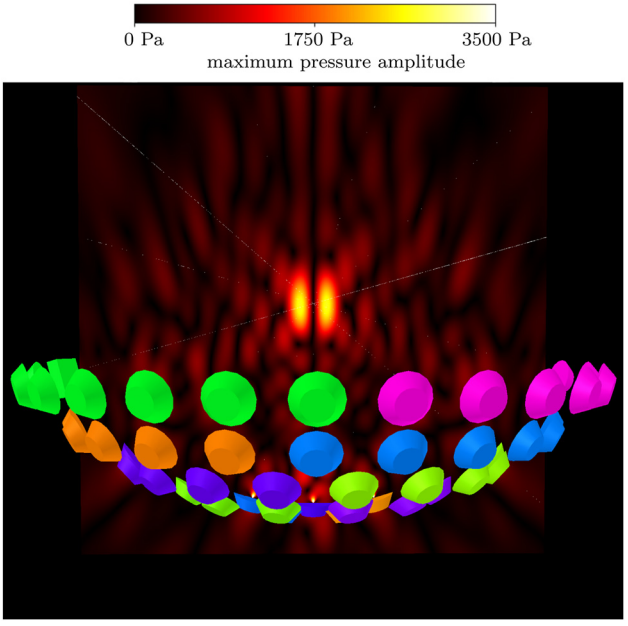

Figure 1: Simulation of the pressure field of a twin trap created with the Ultraino software [20]. Transducers of the same color emit the ultrasonic waves in phase.

2.2 Schlieren photography

Schlieren photography is an optical measurement technique in fluid mechanics. Schlieren are areas in a transparent medium, such as air, that refract light differently than their surroundings [26]. Since the pressure gradients of ultrasonic waves create Schlieren, which are invisible to the naked eye, Schlieren photography can be leveraged to visualize them. A schematic illustration of a Schlieren photography setup is shown in Figure 2 [26]. A light-emitting diode (LED) coupled with a pinhole creates a point light source which is located twice the focal distance ($2f$) of a parabolic mirror. For imaging, a camera is used with a blade placed in front of it. The blade is also $2f$ away from the mirror.

2.3 Imaging concept

In order to facilitate biodiversity research, it is necessary to employ high-resolution imaging techniques with EDOF. This can be achieved by taking a focus series from every viewing angle and focus-stacking. For proper photogrammetric lighting conditions, daylight white light with diffusers are well-suited. To obtain images from multiple perspectives, the invertebrate is rotated by changing the acoustic pressure field while the camera remains stationary. Since the camera cannot enter the acoustic pressure field, a lens with

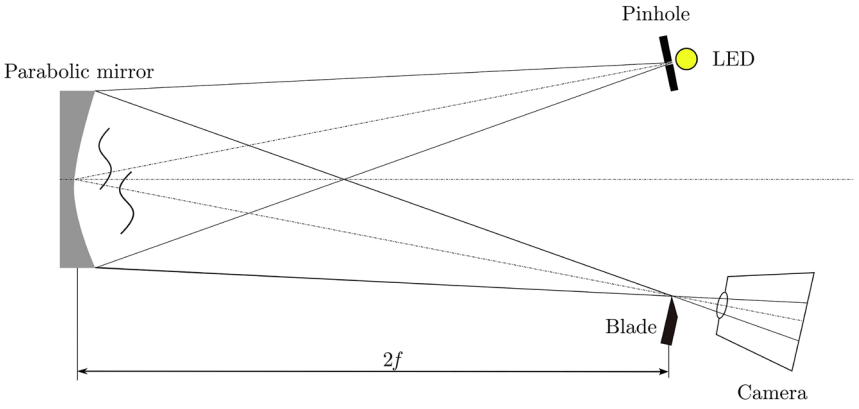

Figure 2: Schlieren imaging setup to allow visualization of the pressure gradients of an acoustic levitator. Image modified from [26].

a long focal length is required to have the insect fill most of the image frame.

3 Results

3.1 Automated close-range imaging

We present an automated system combining photogrammetry with acoustic levitation for the multi-perspective imaging of small insects, shown in Figure 3.

The system consists of a levitation unit using a FPGA combined with an amplifier board to control the 87 ultrasonic transmitters. With this control unit, the insect can be rotated to take EDOF images with a stationary camera with a macro lens. A graphical user interface (GUI) is implemented to facilitate the usage of the system.

3.2 Levitation unit

The levitation unit consists of a half sphere, fitting 87 ultrasonic transmitters. The transmitters are the MA40S4R/S from Murata with a frequency of 40 kHz and a sound pressure level of 120 dB. To ensure a secure fitting of the ultrasound transducers in the half spheres 3D printed sleeves with O-ring seals are used for an elastic fit. The parts are depicted in Figure 4. This also has the advantage that the transducers can easily be changed in case one is damaged.

Figure 3: Photogrammetric imaging system using acoustic levitation. The levitator consists of (1) a half sphere with ultrasonic transducers and is controlled by (2) an FPGA with an amplifier board in the base of the setup. For the image acquisition, (3) a macro camera standing on a tripod is used. For a suitable lighting condition (4) a custom made light LED ring with diffusors is used.

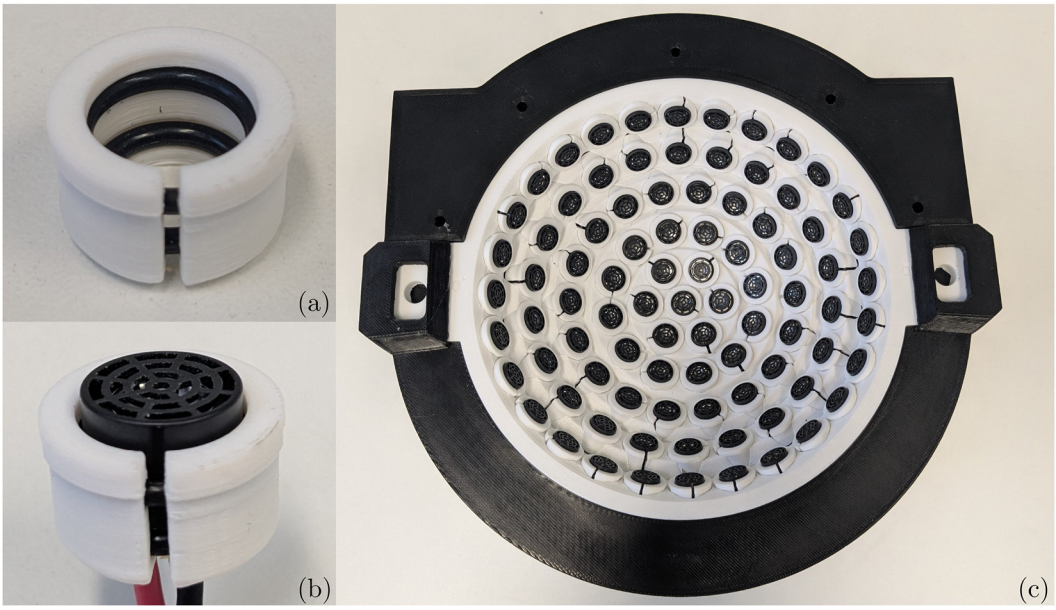

Figure 4: The transducer sleeve and the casing. (a) The transducer sleeve is a 3D printed part with two O-ring seals. (b) The transducer is in the sleeve. (c) The transducer casing with mounted transducers.

To connect the transducer sleeves to the casing, they are simply pushed into the designated holes of the casing. The transducers are arranged in a ring shape on the half sphere, with the transmitter number increasing in every ring. The first ring has one transmitter, the second six, the third 12, the fourth 18, the fifth 22, and lastly the sixth 28 creating a total of 87 transmitters on the half sphere. The transmitters have a small diameter of only 10 mm, which allows the high number of ultrasonic transmitters to be installed in the compact space. This also increases the number of possible degrees of freedom for generating the acoustic field, and smaller rotation angle increments are possible.

Since 3D reconstruction requires images from multiple perspectives, a fine-resolution rotation increment by the ultrasonic transducers is aimed for. Initial attempts to use an Arduino Mega microcontroller as a signal generator as a control unit were unsatisfactory. Only a phase resolution of $\pi/5$ could be achieved, and due to the hardware-related phase shift, the acoustic field required for rotation via acoustic levitation could not be generated. For this reason, an FPGA board (Alchitry AU) is used as a signal generator for the levitator [27]. This enables a much finer phase resolution of $\pi/1250$. The FPGA board has 102 IO pins. We use 87 pins to control the 87 transducers individually. It is possible to use further nine pins for controlling nine more transducers. The final six pins can be used to expand the system and add a second board via the remaining pins to control more transducers. The logic signals of the FPGA are amplified by metal–oxide–semiconductor field-effect transistors (MOSFETs) gate drivers in an amplifier stage from 3.3 V to a maximum of 30.0 V. The FPGA and amplifier board are depicted in Figure 5. They are connected with the four onboard GPIO connectors of the FPGA board. Using such a signal generation allows rotating the specimen by 5° angles. Two movies of a rotating insect can be found in the Supplementary Material.

The positioning accuracy of the levitated specimen is primarily governed by the stability of the acoustic pressure field and the phase resolution of $\pi/1250$ of the FPGA board. Small positional oscillations around the equilibrium point are present, as is typical for acoustic levitation systems, but the impact on the image acquisition is negligible due to the short exposure times. The influence of these positioning uncertainties on subsequent potential 3D reconstruction is mitigated by using structure from motion (SFM), which estimates and optimizes camera poses.

No visible deformation of delicate structures such as wings or antennae was observed during levitation and rotation of the specimen. Across different rotation angles, the specimen appeared to maintain a consistent morphology. However, subtle displacements cannot be excluded and were not quantitatively assessed in this study.

A schematic overview of the electronic components of the system is depicted in Figure 6. The phases of the individual transducers, which make up the stationary pressure field, are stored in frames. These are calculated with the open-source Ultraino software [20]. An animation is a collection of such frames. Changing the frames in an animation

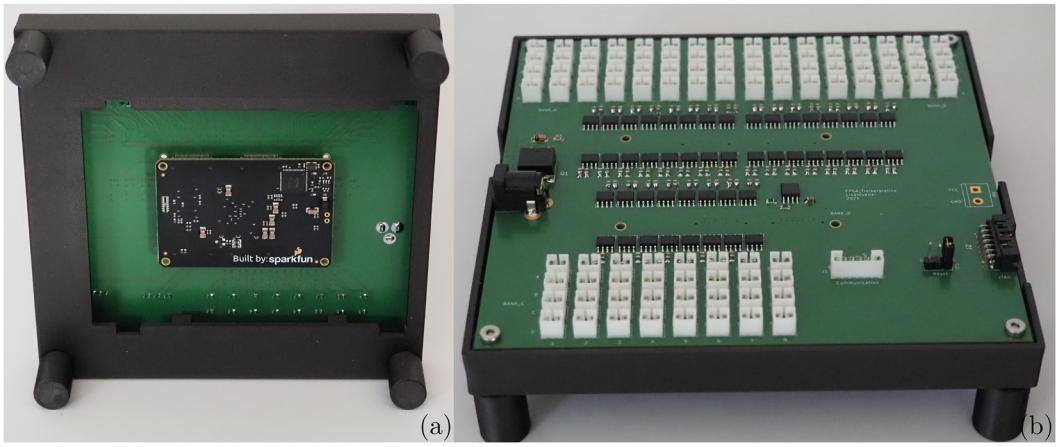

Figure 5: The FPGA and amplifier board. (a) The FPGA board is located under the amplifier board. (b) The amplifier board with a maximum of 96 outputs for the ultrasonic transmitters.

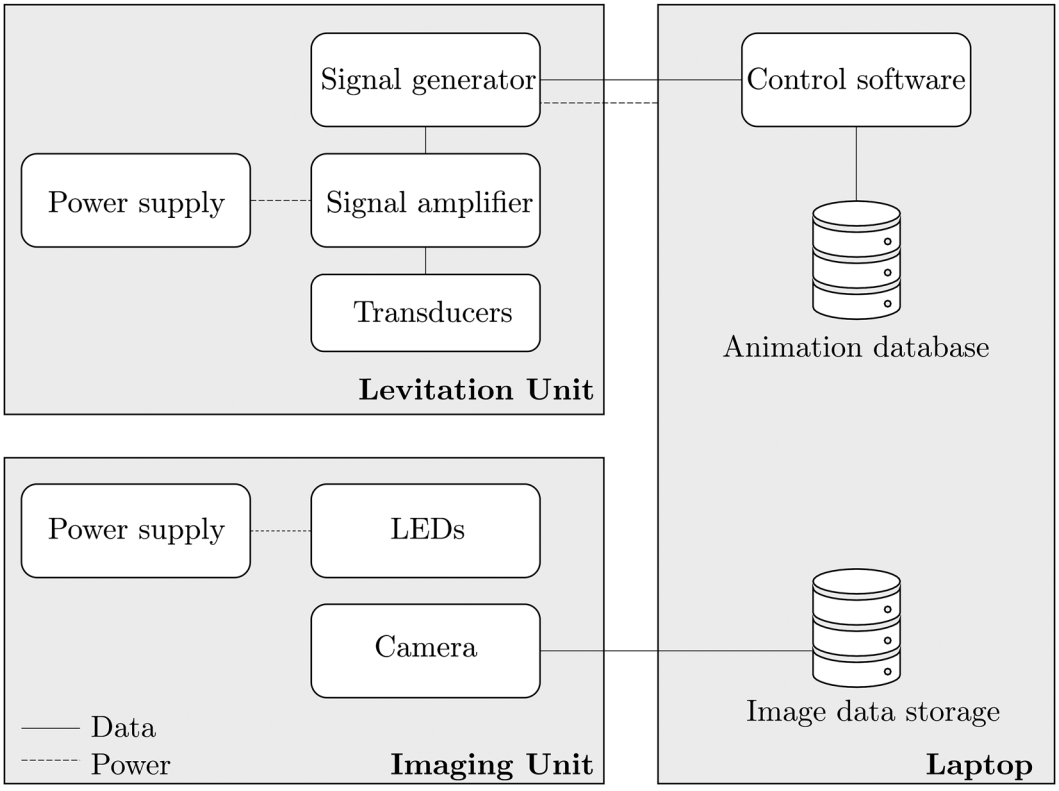

Figure 6: A schematic overview of the electronic components. The control software receives the animations for the levitation unit from the animation database. The generated signal is amplified by the amplifier board, which is powered by an external power supply, and passed on the transducers. The LEDs are also powered by an external power supply and the camera stores the data onto the laptop. The lines depict the data transfer while the dotted lines represent the power supplies.

causes the pressure field to change, allowing a rotation of the levitated object. The animations are passed on to the control software, which transfers the data to the signal generator. The control software sets the outputted frame and the amplitude of the sound wave, which is realized with a pulse-width modulation of the transducer signal.

A finite state machine (FSM) is implemented on the FPGA to control the output signals to the transducers and

store phase information in Random-Access Memory (RAM). The phase information are values between 0 and 2500 and indicate the start of the duty cycle. Additionally, two output buffers are implemented to ensure the signal output to the transducers is not interrupted. When the frame is changed, the new frame, which should be output next, is loaded into the inactive buffer while the old frame remains active. Only after this step is the new frame output.

A signal generator module is implemented for each transducer. It uses the phase information from the active output buffer and the set duty cycle to generate the 40 kHz output signal for the respective transducer. For this purpose, the 100 MHz clock of the signal generator is used, which clocks a common counter for all transducers that counts up to 2500 to create the 40 kHz signal. Each signal generator compares the current count value with the phase information and duty cycle from the output buffer. The signal generator is powered by the connected laptop and passes the signal for each transducer via the connection GPIO pins to the amplifier stage, which is powered by an external power source. The LEDs in the imaging unit are also connected to an external power supply and the camera, which is connected to the laptop via USB, saves the images onto the laptop for further processing.

3.3 Visualization of the pressure fields

As the pressure gradients of the ultrasonic waves are invisible to the naked eye, Schlieren photography is used to visualize them and to verify a rotation of the field. The images are taken with an off-axis arrangement using a parabolic mirror with a diameter of 150 mm and a focal length f = 550 mm, and a point-shaped LED as a light source, as depicted in Figure 2. The resulting Schlieren photos and the simulated pressure fields at 0° and 90° are shown in Figure 7.

The Schlieren photos do not directly depict the pressure field but rather the pressure gradient. Therefore, changes in the pressure field are seen as the light and dark areas in the Schlieren photos. Furthermore, the changes in the pressure field are only shown in one direction depending on the blade position of the Schlieren setup. Rotating the pressure field and the blade by 90° is predicted to result in an image turned by 90°, which is achieved. In Figure 7, the images (c) and (d) are the same but rotated by 90°.

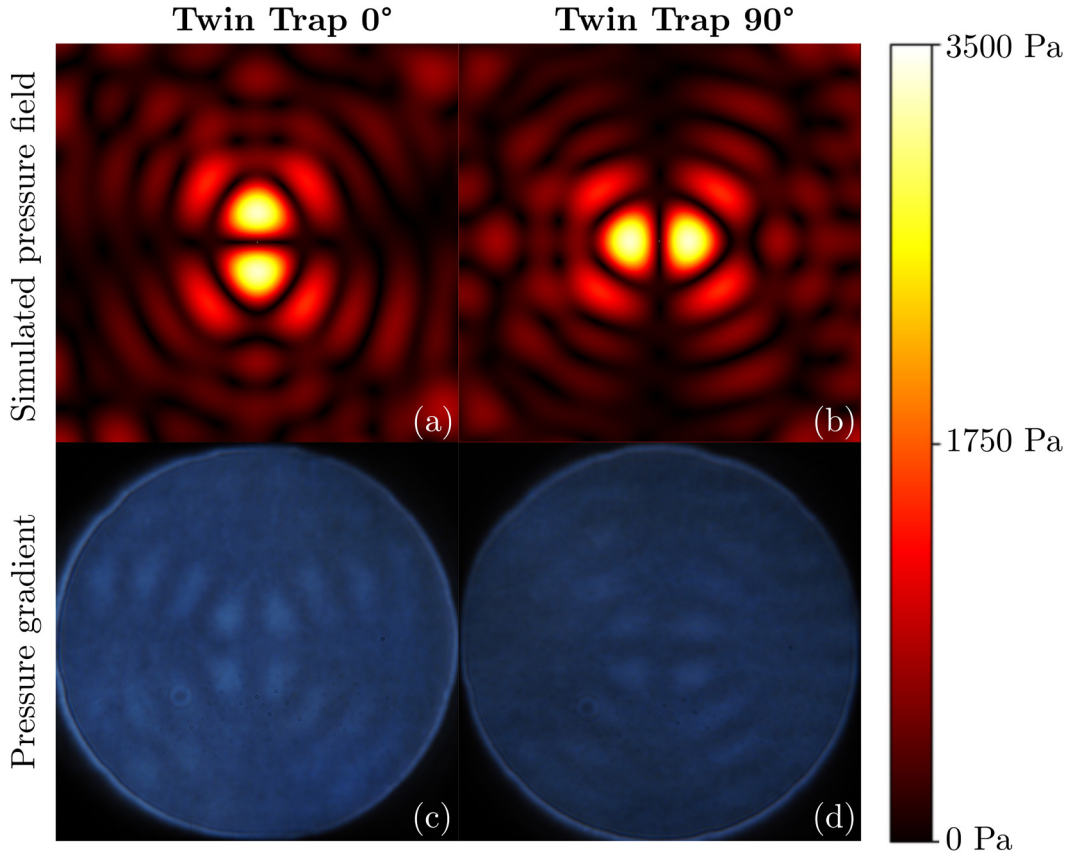

Figure 7: Pressure and pressure gradients of setup with rotating pressure field. The simulated pressure field is calculated with the Ultraino [20] software at (a) 0° and (b) 90°. The Schlieren photos show the pressure gradient (c) in the horizontal direction at 0° and (d) in the vertical direction at 90°. The light and dark areas (c–d) show the changes in the pressure gradient.

3.4 Imaging unit

The imaging unit comprises an OM-D E-M1 Mark III camera (Olympus) with a M. Zuiko Digital ED 90 mm F3.5 Macro (Olympus) lens, which can be seen in Figure 3. The lens's long focal length allows imaging of small specimens from a larger distance (>60 mm) to ensure the camera does not enter the ultrasonic field, which would disrupt the levitation. An M. Zuiko Digital 2× Teleconverter MC-20 (Olympus) can be placed in-line to increase the magnification further, depending on the size of the invertebrate. To establish a suitable lighting condition, a custom made light ring is used. It uses 5500 K LED strips with diffusers, and is attached to a 3D printed white background.

3.5 Workflow

The workflow of the multi-perspective imaging is depicted in Figure 8. First, the system is initialized using the GUI. In the GUI, the user can specify the angle of rotation and the time between the rotation steps. The specimen is taken out of the ethanol it is conserved in. It is then placed on a paper towel for a few minutes to ensure there are no big weight fluctuations during the imaging process. The insect is then placed in the acoustic pressure field with a thin organza mesh, which is only used for positioning and removed before imaging. Schlieren imaging confirmed that the thin mesh does not affect the acoustic field during positioning. After positioning, the mesh is removed.

A focus bracketing series with 40 images is then taken with the camera. We opted for a focus bracketing series with more images than needed, to ensure the camera does not have to be readjusted when the insect is rotated and taking 40 images only requires 3 s. The camera is controlled by a laptop connected via a USB-C cable to avoid motion by pressing the trigger. The insect is then rotated by 5° around the vertical axis by the levitator to the next position, and a new focus bracketing series is taken. This is repeated until one full circle with 72 perspectives is imaged. Focus stacking is performed with Helicon Focus software (HeliconSoft) and stored on the connected laptop.

To ensure genome sequencing is still possible on the insects after drying and imaging them, they are placed back into the ethanol after the imaging process. Once placed back into the ethanol they sink to the bottom of the vial and do not float, which facilitates genome sequencing in an ensuing step.

3.6 Resulting images

In our experiments, winged insects up to 25 mm could be levitated and rotated successfully regardless of their morphology. The same animation created in the Ultraino software [20] is used, simplifying the process since the pressure fields do not have to be generated and calculated for each individual insect. Sample images taken with the proposed imaging system are depicted in Figure 9. Further viewing angles are provided in the Supplementary Material. The images show insects of differing sizes and morphologies rotated by 90°. The images are typically taken with an F/10 aperture, an ISO of 1600, and a shutter speed of 1/1600 s. The fast shutter speed is used to minimize motion blur caused by small oscillations of the levitated specimen. The high ISO value compensates for the reduced exposure time by maintaining sufficient image brightness. Imaging each perspective with 40 images per stack takes 3 s, and imaging 72 perspectives in total takes under 10 minutes, including the rotation time of the insect and the time it takes for it to settle in the new pressure field.

3.7 3D reconstruction as a potential use case

The multi-perspective images can be used for various use cases. These include digitization, taxonomic work, and morphological work. The multi-perspective images can be

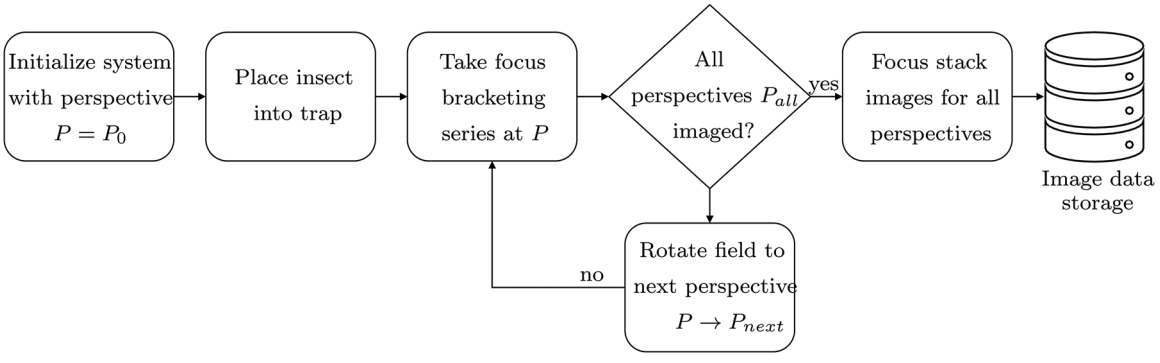

Figure 8: Multi-perspective imaging using acoustic levitation workflow. After initialising the system, the insect is placed into the acoustic pressure field. A focus bracketing series is taken at each desired perspective. Then the images are stacked and stored on the connected laptop.

Figure 9: Images with 90° rotation taken with the proposed imaging system.

stored as 2D images, as a partial 3D reconstruction, or as a full 3D reconstruction. For taxonomic work, the multiple perspectives can help researchers better distinguish closely related species by offering more viewpoints. Furthermore, this provides additional training data for machine learning classification. The images and 3D models can also fully describe the anatomy of newly discovered insects.

To show the resulting images taken with the levitation setup can be used for these use cases a partial and a full reconstruction is performed. For the partial 3D reconstruction the RGB images taken with this system are converted into RGB-D images. RGB-D images combine standard RGB color information with per-pixel depth values, enabling a 3D representation of the scene. This has the advantage of creating new views out of a single perspective. To obtain the missing depth information the Depth Anything V2 model is used [28]. The model has an accuracy of over 95 % on their proposed benchmark. However, the generated depth maps represent learned estimates rather than physically measured depth values and may therefore introduce local inaccuracies. The combined representation can be depicted in a 3D space, as shown in Figure 10.

For a full 3D reconstruction, NeRF is used. Compared to traditional 3D modeling approaches, which create a 3D mesh, NeRF creates a 3D representation of a scene and can create novel views. It represents a scene as a continuous function that maps spatial position and viewing direction to density and color. The expected color of a pixel C(**r**) is obtained by integrating radiance along a camera ray **r**(t) [29]:

$$C(\mathbf{r}) = \int_{t_n}^{t_f} T(t)\sigma(\mathbf{r}(t))\mathbf{c}(\mathbf{r}(t), \mathbf{d})\,\mathrm{d}t \tag{5}$$

$$T(t) = \exp\left(-\int_{t_n}^{t} \sigma(\mathbf{r}(s))\mathrm{d}s\right) \tag{6}$$

$$\mathbf{r}(t) = \mathbf{o} + t\mathbf{d} \tag{7}$$

where σ is the volume density and t_n and t_f are the near and far bounds.

If needed, the 3D reconstruction can be converted into a 3D mesh to have both file formats. Furthermore, they can create a 3D representation using less images and can handle transparent structures, such as the wings of an insect.

An example of a reconstruction using NeRF with the proposed imaging system is shown in Figure 11. The stacked images are first segmented from the background using the machine learning model Grounded SAM [30]. This ensures fine structures, such as the antennae, are not connected in the resulting 3D reconstruction. The camera poses of the segmented images are matched using Colmap [31], [32] and

Figure 10: Creation of a partial 3D point cloud representation using depth maps. (a) Depth map created with depth anything V2. (b) Point cloud of the RGBD image.

Figure 11: 3D reconstruction using Colmap and NeRF.

transferred to Nvidia's NGP Instant NeRF [33]. The total time to match the poses and create the model was under an hour on a PC equipped with a NVIDIA GeForce RTX 3080 graphics card.

4 Summary and discussion

The proposed system for automated close-range imaging using acoustic levitation allows multi-perspective imaging of small, winged insects up to 25 mm by levitating and rotating the specimens in an acoustic pressure field. The signal sent to the ultrasonic transmitters is generated using a FPGA and an amplifier board. The rotation of the pressure field is qualitatively examined using a Schlieren setup prior to imaging the insects. A macro camera and suitable lighting conditions enable taking high-quality EDOF images of the levitated specimen. By using many 3D printed parts and many standard components, the overall cost of the system is sufficiently low that it could be implemented in biodiverse, tropical countries [9].

Unlike other photogrammetric imaging systems used in biodiversity research, the object is not manipulated using an insect pin or other physical attachments. This ensures the specimen is not physically damaged, and since the imaging is sufficiently quick in comparison to previous systems, the specimens do not dry to such an extent that they float on ethanol after returning them to a vial. This means that further DNA barcoding is facilitated. Furthermore, all parts of the specimen can be viewed as no pin is blocking the view of the specimen. It also ensures insects too small to be pinned can be imaged, and the time and labor-intensive pinning process is avoided.

Using a twin trap as a pressure field allows not only the levitation and imaging of very small insects but also of insects larger than 4 mm, which offers a versatile use of the proposed system. Unlike previous levitation systems, our system can rotate levitating objects without any prior knowledge of their geometry.

In a future system, the ultrasonic transmitters could be controlled in such a way that rotation in several axes would be possible, thus providing further perspectives of the specimen and improving the quality of the 3D representation. Alternatively, a few more perspectives can be captured by slightly tilting the camera to capture perspectives above and below the insect. More automation steps, including automated sample insertion into the acoustic field, would

further improve the system and reduce human error and manual work. Currently, our system requires less than 10 minutes for imaging a full rotation with which it is possible to achieve a 3D reconstruction, as opposed to other systems requiring at least 40 min for the data acquisition process [3].

Using the multi-perspective images from this device or the 3D reconstruction, taxonomists can better describe and identify insect species. Body parts, which may be hidden in an image from a single perspective, are made visible through multiple angles. Furthermore, morphometric measurements of insects are more precise using a 3D reconstruction than a 2D image [3]. All this will greatly aid in collecting trait information on insect species [12]. The latter is important for assessing the relative importance of insect species within an insect community and thus aids in assessing the health of ecosystems.

It should be noted that the reconstruction methods used in this work's pipeline do not inherently produce metrically accurate 3D reconstructions. Instead, they optimize for photometric consistency across views, which can result in geometrically plausible but not strictly accurate representations. The reconstructed models are suitable for visualization, qualitative analysis, and machine learning applications; however, their use for precise physical measurements requires additional validation. Future work will therefore include benchmarking against established measurement techniques to quantify geometric accuracy and investigate how the multi-perspective imaging reduces parallax-related errors compared to single-view approaches.

Furthermore, additional NeRF variants will be investigated in future work. For instance, Fourier feature mapping may improve reconstruction quality by enabling better representation of high-frequency details [34]. Additionally, Bundle-Adjusting Neural Radiance Fields (BARF) could be employed, which jointly optimizes scene representations and resolves large camera pose misalignment [35].

Although the application of this work focuses on biodiversity research, it can also be applied in other research areas where small, light objects are analyzed. For instance, it could image small seeds, crystals and liquid droplets for analysis.

In summary, the proposed system for automated imaging offers a promising solution for close-range photogrammetry of small, light objects. It is a non-destructive approach to multi-perspective imaging and offers the basis for 3D reconstruction.

Acknowledgments: We want to gratefully thank Daniel Moser for his hardware support, Moritz Lindmüller for his electronics expertise, and Bernhard Schurian for his imaging expertise, all of which were essential to the success of this project.

Research ethics: Not applicable.

Informed consent: Not applicable.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Use of Large Language Models, AI and Machine Learning Tools: During the preparation of this work the authors used ChatGPT in order to improve clarity and fluency during the writing process. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Conflict of interest: The authors state no conflict of interest.

Research funding: Funding was provided by the Center for Integrative Biodiversity Discovery at the Museum für Naturkunde Berlin and Grant #ZF4717901SK9 of the program Natural, Artificial and Cognitive Information Processing (NACIP) of the Helmholtz-Association, Germany.

Data availability: The data can be obtained on request from the corresponding author.

References

[1] C. Villa, "Forensic 3D documentation of skin injuries," *Int. J. Leg. Med.*, vol. 131, no. 3, pp. 751—759, 2017.

[2] V. Evgenikou and A. Georgopoulos, "Investigating 3D reconstruction methods for small artifacts," *Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci.*, vols. XL-5/W4, pp. 101—108, 2015.

[3] B. Ströbel, S. Schmelzle, N. Blüthgen, and M. Heethoff, "An automated device for the digitization and 3D modelling of insects, combining extended-depth-of-field and all-side multi-view imaging," *ZooKeys*, vol. 759, pp. 1—27, 2018.

[4] C. V. Nguyen, D. R. Lovell, M. Adcock, and J. La Salle, "Capturing natural-colour 3D models of insects for species discovery and diagnostics," *PLoS One*, vol. 9, no. 4, 2014, Art. no. e94346.

[5] F. Plum and D. Labonte, "scAnt — An open-source platform for the creation of 3D models of arthropods (and other small objects)," *PeerJ*, vol. 9, 2021, Art. no. e11155.

[6] A. Jankielsohn, "The redesign of sustainable agricultural crop ecosystems by increasing natural ecosystem services provided by insects," *Adv. Ecol. Environ. Res.*, vol. 2, pp. 365—381, 2017.

[7] H. Shirali, *et al*., "Image-based recognition of parasitoid wasps using advanced neural networks," *Invertebr. Systemat.*, vol. 38, no. 6, 2024. https://doi.org/10.1071/is24011.

[8] A. Srivathsan, *et al*., "ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by everyone, for everyone," *BMC Biol.*, vol. 19, no. 1, p. 217, 2021.

[9] M. Brydegaard *et al*., "Towards global insect biomonitoring with frugal methods," *Philos. Trans. R. Soc. B: Biol. Sci.*, vol. 379, no. 1904, 2024, Art. no. 20230103.

[10] H. Tatsuta, K. H. Takahashi, and Y. Sakamaki, "Geometric morphometrics in entomology: Basics and applications," *Entomol. Sci.*, vol. 21, no. 2, pp. 164—184, 2018.

[11] O. Dry, S. Muthu, J. Tong, C. Nguyen, and L. Petersson, "IQScan: Automatic insect loading and 3D scanning," in *2023 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Port Macquarie, Australia, IEEE, 2023, pp. 205—211.
[12] R. Meier, E. Hartop, C. Pylatiuk, and A. Srivathsan, "Towards holistic insect monitoring: Species discovery, description, identification and traits for all insects," *Philos. Trans. R. Soc. B: Biol. Sci.*, vol. 379, no. 1904, 2024, Art. no. 20230120.
[13] L. Wührl, L. Rettenberger, R. Meier, E. Hartop, J. Graf, and C. Pylatiuk, "Entomoscope: An open-source photomicroscope for biodiversity discovery," *IEEE Access*, vol. 12, pp. 11785—11794, 2024.
[14] L. Wührl *et al*., "DiversityScanner: Robotic handling of small invertebrates with machine learning methods," *Mol. Ecol. Resour.*, vol. 22, no. 4, pp. 1626—1638, 2022.
[15] J. Ärje *et al*., "Automatic image-based identification and biomass estimation of invertebrates," *Methods Ecol. Evol.*, vol. 11, no. 8, pp. 922—931, 2020.
[16] L. Wührl, K. Rotmann, R. Meier, E. Hartop, N. Klug, and C. Pylatiuk, "DiversityScanner-360°: An automated system for digitizing invertebrate bulk samples," in *2023 3rd International Conference on Robotics, Automation and Artificial Intelligence (RAAI)*, Singapore, IEEE, 2023, pp. 226—230.
[17] A. Srivathsan, *et al*., "Convergence of dominance and neglect in flying insect diversity," *Nat. Ecol. Evol.*, vol. 7, no. 7, pp. 1012—1021, 2023.
[18] M. A. B. Andrade, N. Pérez, and J. C. Adamowski, "Review of progress in acoustic levitation," *Braz. J. Phys.*, vol. 48, no. 2, pp. 190—213, 2018.
[19] C. J. Benmore and J. K. R. Weber, "Amorphization of molecular liquids of pharmaceutical drugs by acoustic levitation," *Phys. Rev. X*, vol. 1, no. 1, 2011, https://doi.org/10.1103/physrevx.1.011004.
[20] A. Marzo, T. Corkett, and B. W. Drinkwater, "Ultraino: An open phased-array system for narrowband airborne ultrasound transmission," *IEEE Trans. Ultrason. Ferroelectrics Freq. Control*, vol. 65, no. 1, pp. 102—111, 2018.
[21] L. Cox, A. Croxford, B. W. Drinkwater, and A. Marzo, "Acoustic lock: Position and orientation trapping of non-spherical sub-wavelength particles in mid-air using a single-axis acoustic levitator," *Appl. Phys. Lett.*, vol. 113, no. 5, 2018, Art. no. 054101.
[22] P. Helander *et al*., "Omnidirectional microscopy by ultrasonic sample control," *Appl. Phys. Lett.*, vol. 116, no. 19, 2020, Art. no. 194101.
[23] F. Sundblad, D. Iablonskyi, A. Holmstrom, A. Salmi, and E. Haggstrom, "Position and orientation control of complex-shaped samples in acoustic levitation," in *2022 IEEE International Ultrasonics Symposium (IUS)*, Venice, Italy, IEEE, 2022, pp. 1—4.
[24] L. P. Gorkov, "Forces acting on a small particle in an acoustic field within an ideal fluid," *Doklady Akademii Nauk Sssr*, vol. 140, p. 88, 1961.
[25] H. Bruus, "Acoustofluidics 7: The acoustic radiation force on small particles," *Lab Chip*, vol. 12, no. 6, pp. 1014—1021, 2012.
[26] G. S. Settles, *Schlieren and Shadowgraph Techniques*, Berlin, Heidelberg, Springer, 2001.
[27] S. Zehnter and C. Ament, "A modular FPGA-based phased array system for ultrasonic levitation with MATLAB," in *2019 IEEE International Ultrasonics Symposium (IUS)*, Glasgow, United Kingdom, IEEE, 2019, pp. 654—658.
[28] L. Yang, *et al*., "Depth anything V2," *arXiv*, 2024. https://doi.org/10.48550/arXiv.2406.09414.
[29] B. Mildenhall, P. Srinivasan, M. Tancik, J. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *arXiv*, 2020. https://doi.org/10.48550/arXiv.2003.08934.
[30] T. Ren, *et al*., "Grounded SAM: Assembling open-world models for diverse visual tasks," *arXiv*, 2024. https://doi.org/10.48550/arXiv.2401.14159.
[31] J. Schönberger, F. Fraundorfer, and J.-M. Frahm, "Structure-from-motion for MAV image sequence analysis with photogrammetric applications," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences — ISPRS Archives*, vol. 40, 2014, pp. 305—312.
[32] J. L. Schönberger, E. Zheng, J.-M. Frahm, and M. Pollefeys, "Pixelwise view selection for unstructured multi-view stereo," in *Computer Vision — ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, eds., vol. 9907, Cham: Springer International Publishing, 2016, pp. 501—518.
[33] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *ACM Trans. Graphics*, vol. 41, no. 4, pp. 1—15, 2022.
[34] M. Tancik, *et al*., "Fourier features let networks learn high frequency functions in low dimensional domains," *arXiv*, 2020. https://doi.org/10.48550/arXiv.2006.10739.
[35] C.-H. Lin, W.-C. Ma, A. Torralba, and S. Lucey, "Barf: Bundle-adjusting neural radiance fields," *arXiv*, 2021. https://doi.org/10.48550/arXiv.2104.06405.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/auto-2025-0139).

Bionotes

Nathalie Klug
Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344, Baden-Württemberg, Germany
nathalie.klug@kit.edu
https://orcid.org/0009-0007-8185-8779

Nathalie Klug completed her Bachelor's and Master's degree in mechanical engineering at the Karlsruhe Institute of Technology in Karlsruhe, Germany, in 2019 and 2022, respectively. Currently, she is working as a doctoral researcher in the Biomedical Engineering and Robotics group of the Institute for Automation and Applied Informatics at the Karlsruhe Institute of Technology in Eggenstein-Leopoldshafen. Her research focuses on the automated 3D modeling of small objects for biodiversity research.

Markus Kramer
Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344, Baden-Württemberg, Germany

Markus Kramer received his Bachelor's and Master's degree in mechanical engineering from the Karlsruhe Institute of Technology in Karlsruhe, Germany, in 2019 and 2024, respectively. He is currently working as a doctoral researcher at the Institute for Automation and Applied Informatics at the Karlsruhe Institute of Technology, focusing on the development of automated devices for biodiversity research.

Fahri Mazrek
Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344, Baden-Württemberg, Germany

Fahri Mazrek received the Bachelor's degree from the Philipps University, Marburg, Germany, in 2023. Until October 2024 he aided the development of 3D model creation pipelines for species discovery as a working student at the Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology.

Rudolf Meier
Center for Integrative Biodiversity Discovery, Leibniz Institute for Evolution and Biodiversity Science, Museum für Naturkunde, Invalidenstraße 43, 10115, Berlin, Germany and Institute for Biology, Humboldt University Berlin, Unter den Linden 6, 10117, Berlin, Germany

Rudolf Meier received the Ph.D. degree in biology from Cornell University, in 1995. Until 1997, he was with the American Museum of Natural History. From 1997 to 2003, he was an Associate Professor and a Curator with the University of Copenhagen. From 2003 to 2011, he was an Associate Professor. From 2011 to 2021, he was a Professor with the National University of Singapore. Since 2021, he has been the Director of the Center for Integrative Biodiversity Discovery, Museum für Naturkunde, Berlin, Germany, and a Professor at Humboldt University, Berlin.

Christian Pylatiuk
Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344, Baden-Württemberg, Germany

Christian Pylatiuk received the M.D. degree from the University of Marburg, Germany, in 1997. He worked on mechatronic cooperation projects with the University of Heidelberg to develop multifunctional bionic hands and on a gripping orthosis for paraplegics. Since 2014, he has been an Adjunct Professor with the Faculty of Mechanical Engineering and is leading the Research Group "Biomedical Engineering and Robotics," Karlsruhe Institute of Technology (KIT). His research interests include mechatronics, medical technology, lab automation, optical sensor systems, and image analysis.