Grant Agreement Number: 777483 / Acronym: ICEDIG
Call: H2020-INFRADEV-2017-1 / Type of Action: RIA
Start Date: 01 Jan 2018 / Duration: 27 months

REFERENCES:
DELIVERABLE D3.5
Work package WP3 / Task T3.1.2 / Lead: University of Helsinki
Deadline M19

ICEDIG.EU

Innovation and consolidation for large scale digitisation of natural heritage

State of the art and perspectives on mass imaging of pinned insects

Deliverable D3.5

Authors:

Zhengzhe Wu (UH), Anne Koivunen (UH), Hannu Saarenmaa (UH), Myriam van Walsum (PIC), Agnes Wijers (PIC), Luc Willemse (Naturalis), Toivo Ylinampa (NampaWorks Ltd)

Abstract

This report describes the state of art and work in progress on automated methods in mass-digitisation of pinned insects. The report begins by identifying the challenges, which stem from the fact that pinned insects are basically 3D objects and their numbers in collections are huge, up to one billion objects in Europe. A ten-fold increase in the speed of their digitisation from the current state of art is being sought for. Based on recent developments, state-of-the-art in their digitisation is covered. Important new technologies which seem promising in insect digitisation are described.

The report describes the experiments which the ICEDIG project has carried out in order to find new innovations in mass-digitisation of pinned insects. We identify six such possible approaches, scoping their features, applicability, possible benefits and limitations, and make recommendations. Progress on the ongoing developments in three experiments is described in detail.

We conclude that achieving a breakthrough in insect digitisation probably requires a combination of existing and new technologies in novel workflows.

Table of contents

1. Introduction

1.1. Background

The ICEDIG project is challenged by finding the innovations that are needed by the DiSSCo Research Infrastructure to ramp up the pace of digitisation. In its work package WP3, the ICEDIG project is considering how to improve the process of imaging specimens in biological collections. This means physical handling of specimens, i.e., "turning atoms into bits". This is one of the first steps in the overall process of digitisation. Both the specimen and its associated labels need to be imaged. Prior to imaging, each specimen must be tagged with unique identifier, which ideally is machine-readable, such as QR-barcode.

By "digitisation" we mean the extended process that covers tagging with unique identifiers and imaging, but also later steps of data capture from the labels attached to the specime, enhancing the data with georeferencing and by other means, and even extracting DNA and chemicals. Minimum Information for Digital Specimens (MIDS) is a draft specification of different levels of completeness of digitisation (see, e.g., ICEDIG MS35 report, and the forthcoming Deliverable D6.5). MIDS-0 level is the absolute minimum of information worth to share publicly: It requires only a unique identifier, possibly images, and technical metadata of the imaging process. Imaging has to be done in such a way that it facilitates elevating the specimen data to further MIDS levels. This basically requires that the specimen is findable through its technical metadata, and that the specimen be photographed in such detail that it can be distinguished from nearby species (which, however, may not always be possible). Furthermore, out of the labels as many as possible should have been photographed, as well as the unique ID in the barcode.

It is understood that imaging shall be done as fast as possible in an automated way, and data entry should not slow down imaging. Any detailed and time-consuming data entry would follow later, which could easily be years later. Data entry from images is a separate phase: This approach allows extended data entry to be done as separate steps, depending on availability of funds and staff. Through minimal data entry (MIDS-1) the images become increasingly findable. Extended data entry (MIDS-2 and up) will still be needed to make the data useable for most purposes.

One of the Tasks (Subtasks) of WP3 is concerned with pinned insects. It has been described like this:

> *Subtask 3.1.2: Mass-imaging of pinned insects. Mass-digitisation of pinned insects has only been achieved in some pilot projects until now. We will investigate what it will take to scale up available techniques and new technologies for processing thousands of specimens per day in one facility. Various approaches such as conveyor belt-driven automatic imaging, cameras attached to robotic hands, and multispectral imaging will be explored, and when needed, piloted with the assistance of subcontractors who know important new technologies.*

The Deliverable has been defined like this:

> *D3.5 : State of the art and perspectives on mass imaging of pinned insects. A report containing a concise review of methods used to date for mass imaging of insects and of their performance and scalability; descriptions of novel approaches and their potential performance; and, where relevant, descriptions of results of pilot trials of novel techniques.*

The report aims to give recommendations on how to significantly improve the throughput by comparing current practice results with trials and outlook of added (new) technology.

This deliverable report is a further development of the milestone reports MS 11 and MS12, which lay ground to the requirements.

1.2. Objectives of the study

According to the Description of the Action, the ICEDIG project is looking for solutions which would allow digitising a significant part (such as 50%) of important public collections in foreseeable time (such as 25 years). This would require a digitisation capacity ten times faster and cheaper than what exists today. While a working solution to digitise 2D objects such as herbarium sheets does exist (see D3.6), digitising insect collections is the paramount challenge.

More than one half of all specimens in scientific collections are pinned insects. In Europe this means up to one billion such specimens. Today's fastest mass-digitisation (i.e., imaging) systems for pinned insects can achieve one thousand specimens/ working day but their handling is tedious, requires precision, and operator shifts every 2 hours are advisable (Tegelberg & al 2014, 2017; Price et al 2018).

Why is a ten times faster digitisation rate necessary? A simple computation using the powers of ten shows this:

Number of pinned insect specimens in European public collections: 10^9
Maximum rate of digitisation using the fastest currently available line: 10^3
à Duration: 10^6 days, which equals 2,740 years

Number of DiSSCo institutions with insect collections, which can be equipped with a fast digitisation line: 10^2
à Duration when a parallel infrastructure of one hundred digitisation lines is in place: 10^4 days, which equals 27.4 years

When ten-fold increase in digitisation speed is available: 10^3 days, which equals 2.74 years

These figures are just abstractions, indicating the magnitude. In reality, we must count for some reduction from the above figures. Nevertheless, it seems to be possible to digitise one billion insect specimens in the planned lifetime of DiSSCo (25 years), but only if we can achieve a ten-fold increase over the current state-of-art. This will require installing one hundred digitisation lines in DiSSCo institutions and facilities!

How does this compare against what is already available for herbarium sheets? The state of art (cf. Oever & Gofferje 2014, cf. ICEDIG D3.1 report) is 5,000 specimens /working day, operated by two people but only with considerable logistical support by more staff.

The slowness of imaging pinned insects follows from the fact that they are essentially 3D objects. Although butterflies/moths, dragonflies and similar large-winged insects can be prepared (spread) as 2-dimensional (2D) objects, the fact that the labels are pinned under the insect specimen makes even these samples 3-dimensional (3D). On the other hand, these large-winged insects easily obscure all labels, necessitating removal or spacing of labels in any approach. For instance, among

the 10 million insect specimens held by the Finnish Museum of Natural History, the proportion of specimens belonging in these large-winged groups is about 30%. Furthermore, in other orders of insects, e.g., beetles, bugs, wasps, grasshoppers, there also are substantial proportions of large species which can block viewing of the labels.

In imaging, the labels are often removed manually, which slows down the imaging process. If the need for manual handling of the labels can be skipped, we can easily multiply the imaging speed. There still remains the need to attach a unique identifier in the sample (see the Discussion Section below). So the first question to ask is, how can we avoid handling the labels?

This document outlines several possibilities for achieving this. We first review the state-of-the-art, and their discussion of future potential. We then discuss promising technological advances, such as conveyors, robotics, machine vision, multispectral scanning, 3D modelling in large scale, and possibilities of their integration. Some of these new technologies have not yet been tried for insect digitisation. In a closely related report of the ICEDIG project (Nieva et al. 2018), 3D techniques have been assessed in detail, determining the state of the art of the technologies, workflows, collection types, and existing efforts, and available commercial actors.

The answer may be that we would not use just one approach to digitise all insect collections, but choose the method optimally based on criteria such as wing size and number of labels. This approach was used by Hereld et al. (2017) who classified (a sample of) the Chicago Field Museum insect collection by the physical characteristics of the specimens, and then recommended varying approaches that best fit each specimen type. Some of the best practises that will turn up may consist of using a specific imaging protocol in combination with a data entry protocol, studied by the ICEDIG Work Package 4. Although imaging protocols and data entry protocols can be carried out completely independently, there may be specific combinations of imaging protocols and data entry protocols that can be more efficiently combined than others.

In Section 4, we define a small number of potential experiments of the most promising technologies. Three of these have been turned into prototypes by the ICEDIG project, which is reported in Section 5. When tested in real production, successful prototypes may finally form the basis of operational systems in the DiSSCo infrastructure.

2. State of the art in insect mass-digitisation

Here we briefly describe the state-of-the-art in insect mass-digitisation from the perspective to identify additional, most promising approaches that have not yet been tested fully. There have been a number of earlier reviews on the subject (e.g., Häuser et al. 2005; Blagoderov & Smith 2012; Holovanchov et al. 2014; Brecko & Mathys 2016) and we do not intend to repeat their analyses. Instead, we identify from the discussions and conclusions of earlier studies the way forward for actual new experiments.

- Manual digi-streets. Most insect mass-digitisation still happens manually. Workers enter metadata from specimen labels and often also take pictures. When a number of workers

perform this in an organised way, we talk of "digi-streets". Their performance can be quite fast, for instance, in 2014, the 45,000 specimens in the bumblebee collection of the Smithsonian Institution were imaged in just 40 days (Kutner 2014). The number of workers employed was not reported, though. At Naturalis, data entry of 850,000 specimens has been done using a similar approach. Fine-tuned, digi-streets can be quite effective, but their performance is linearly dependent on the number of human workers employed. Typical data entry pace is 200 specimens/day and photographing pace 70 specimens/day (cf. Heerlien et al. 2015).

- Angled photography. Recently, the Angled Label Image Capture and Extraction system (ALICE) was developed by the Natural History Museum in London, to extract the top-most label information from angled images without removal of the labels from the specimen pin (Price et al. 2018). It was estimated that the base digitisation rate (calculated as (MinRate + 4*MedianRate + MaxRate) /6 from three digitisers) to be 140 and 194 /person/h for specimens, respectively without and with data matrix barcodes.

- Whole-drawer scanning. Reviewed by Holovanchov et al. (2014), this approach takes images of whole drawers of insects, consisting perhaps of hundreds of specimens. The five systems that have been described include GigaPan, GigaPanMicro, Sat-Scan, DScan (Schmidt et al. 2012), and use of a high-resolution medium format camera such as Hasselblad. The camera is either fixed, moved on a fixed motorised mount, or moves on rails. The output is one huge-resolution image of the entire drawer. Supporting software such as the open-source Inselect package can then be used to crop and segment the images so that pictures of individual insects can be produced (Hudson et al. 2015). This approach is very effective and used in a number of museums around the world. The drawback is that no images of the labels will be produced. However, augmenting this method with camera tilting for label capture might offer a wider range of applications. An open call for proposals to demonstrate such system, with a $1 million award for the winning bid, was launched by the "Beyond-the-Box" project in 2015 https://beyondthebox.aibs.org/, but received no entries!

- Conveyor-driven imaging. Despite the success of conveyor-driven imaging in plant imaging since 2008, only one such system has been developed for insects (Tegelberg et al. 2014, 2017). Individual insects that have been mounted on specific pallets are carried into an imaging station, where they are automatically photographed from different angles using up to three DSLR cameras. If labels cannot be seen this way, they must be manually detached and placed on the pallet. Maximum performance of two operators has been 500 specimens/day when handling only one label.

- 3D mass digitisation. 3D imaging is currently being studied by several research groups. Systems like ZooSphere (Kroupa et al. 2014) have achieved 360 degree "high precision" viewing of the specimen, but doing that in massive scale is another matter. ZooSphere does not produce a separate 3D model of the object. Adcock et al. (2014) describe a 3D modelling device which rotates the insect on a turntable while taking pictures. Ströbel et al. (2018) describe an automated device for the digitisation and 3D modelling of insects, combining extended-depth-of-field and all-side multi-view imaging. Their methods produce detailed 3D models of insects, but the system is not aimed for mass-production and no figures of the performance are given. Developments at Argonne National Laboratory (Hereld et al. 2017

and 2018) and at Darmstadt Technological University (Ritz et al. 2018) are underway, and are looking into producing 3D models which would then be rendered with images.

3. Promising new technologies

- Robotics. There still is excessive need for human operators in insect digitisation, in particular when labels need to be handled. This is the case even in feeding insect samples into an automated conveyor line. There are several areas where robotics could potentially help in insect digitisation. These include handling of labels, moving cameras in unobstructed view positions, and transporting drawers and units. There may be other opportunities as well.
    - Handling of labels (removing and reattaching) is high-precision work on delicate objects and therefore slow, requiring practice and careful hands. Handling of insects is a bit less demanding, but still a job for a professional. Handling of units and drawers can be performed even by an inexperienced worker. Can any of this work be performed by a robot? There are high-precision robots of suitable size available on the market, which are already being used for medical and other demanding tasks. They could potentially be used for handling insects, but the difficulty rises from having proper 3D information of the exact positions of the specimens and labels, and then steering the robot movements accordingly. Difficulties also arise from the storage structure of the collections; different trays, units, drawers, etc.
    - A related, but much less demanding job is to move a small camera in proper viewing position. Handling of insects would be avoided, but still there must be very accurate information of the position of the specimens.
    - Handling of units and drawers is an easier job, but the benefits would require that the entire collection is turned into an automated warehouse. Currently, collection cabinets and drawers have been designed for human operators. Letting a human-sized robot handle them would require redesign of both. If this is possible, moving materials in and out of the collection could yield significant benefits. (This is being addressed by ICEDIG Task T3.3.)
- Machine vision and automatic image analysis have penetrated the society in a big way in recent years. This is most notable in traffic, where speed traps, police cars, and road toll stations already scan the register plates of vehicles in real time. Autonomous vehicles are being tested in real situations. Lane-assist is commonplace. Extending the use of these technologies into digitisation is an obvious step. Labels could be automatically extracted from images of pinned insects, corrected for position and angle and then automatically transcribed. (OCR is being addressed by ICEDIG Task T4.1. See also Hudson et al. 2015, Agarwal et al. 2018, Price et al. 2018.) Also, identification of the species through image analysis has shown to be possible in some cases (Valan et al. 2019).

- 3D modelling, LiDAR. Building of 3D models of individual insects and unit trays / drawers may be needed for two reasons: proper positioning of robot arms and having a digitised model of the insect itself. Both can be approached by building a 3D model of the target. Building a 3D model is different from just photography of a 3D object – there will be a digital object with coordinates in 3D. For controlling the movement of a robot arm, rendering the surface of the object is not needed, but when digitising the actual object, rendering is very much necessary. LiDAR technology (Light Detection and Ranging) is based on laser beams and is widely used in landscape-scale digitisation of terrain and vegetation. It can also be used in small scale, although LiDAR is meant for larger objects. Its accuracy is currently not high enough for insect scale. Laser triangulation, however, can reach spacing on the micron level, so that is relevant. Further details are provided in the forthcoming ICEDIG deliverable D3.7 on 3D modelling *(Rapid 3D capture methods in biological collections and related fields)*. Also see the related Wikipedia article[1].

- Terahertz, time-gated, multispectral imaging. According to Redo-Sanchez et al. (2015), terahertz time-domain spectroscopy (THz-TDS) is a leading method for spectroscopy, imaging and non-destructive testing in the frequency range of 0.1–10 THz. The method can detect structural defects in foams, wooden objects, plastic components, composites, pharmaceutical products' coatings and cultural artefacts. In contrast to infrared-based time-of-flight cameras, optical coherent tomographic techniques and X-ray techniques, THz-TDS provides both fine time resolution and broadband spectral signatures for a variety of dielectric materials. These advantages have motivated researchers to use computational techniques to empower the yet-maturing THz hardware. Despite the prevalence of sub-millimetre layered structures in industry, biology and objects of cultural value, conventional THz-TDS is incapable of deep content extraction for three well-known reasons: signal-to-noise ratio (SNR) drops with depth (or increasing number of layers), the contrast of the content is much lower than the contrast between dielectric layers, the content from deeper layers are occluded by the content from front layers. Therefore, Redo-Sanchez et al. (2015) introduce a time-gated spectral imaging technique that overcomes all of these challenges to extract occluding content from layers whose thicknesses and separations are comparable to the wavelength.

4. Experiments considered by ICEDIG

Below we describe a number of potential tests with various new approaches. The tests described are not all similar, and some derive from other tests and combine various technologies. They deal with imaging specimens and unit trays, static setups, and conveyor belt-driven approaches. We assume that achieving a quantum leap in insect digitisation probably requires a combination of various advanced technologies, such as conveyors and 3D photogrammetry, and most tests envisioned below follow this approach.

[1] https://en.wikipedia.org/wiki/3D_scanning#Strengths_and_weaknesses

4.1. Accept minimal label information

a) Features

In this approach we place individual specimens in an imaging station (manually or by conveyor) and image the specimens without removing the labels. One shot will be made from above and another from a 30-degree angle from the side. This allows capturing the topmost label. If the labels are spaced out well, maybe also other labels can be captured, as well.

b) Applicability

This approach has been used in operational scale while imaging the entire Coleoptera collection of Gunnar Blomqvist at Digitarium (Tegelberg et al. 2014). Using conveyor-driven imaging, a total of 12,400 specimens of all sizes (representing the entire beetle fauna in Finland) were processed in 50 days, giving a rate of 248 specimens /day. This rate is rather slow, since the workers were not experienced entomology curators, and time was spent spacing out the labels for optimal viewing. Also time was spent in reorganising the collection from original boxes to unit trays.

This approach can be applicable in situations where the collection is rather uniform and there is only little information in labels, such as collector's field number, and there are not too many labels. This approach will also work best for other than large-winged insects, which actually constitute 70% of all specimens.

c) Expected benefits

2-3 fold speed increase compared to the basic practice of removing and reattaching labels (Tegelberg et al. 2014).

d) Difficulties and limitations

Labels below the topmost will not necessarily be imaged. How much data will be lacking because of this depends on the collection. If the top label contains all the essential information (such as collector name and a field number), it may just be enough. Furthermore, additional data capture from labels attached to the drawers or unit trays (i.e., MIDS-1 level data) done on the side of imaging, such as entering the taxon name and major geographic area, may supplement the imaging process so that this approach is worthwhile.

e) Recommendation

We already know how to do this, so there is no need for further tests. The question is whether it is worthwhile to obtain such a limited data. When optimising the costs of the total digitisation effort, this approach might have significant role in digitising many collections of certain kinds of specimens at low cost. So this approach should be taken into account in final cost books. Putting that in more general way, this implies that we have to describe collections that are fit to be tackled by this process. One of the conditions would be that the technique is ideal for insects with only one label.

4.2. Multiple webcams

a) Features

This approach is similar to the previous, but adds a number of small webcams for capturing the labels from a number of different angles and directions. It would mean taking one image from above of the specimen and as many as ten images of the labels. If the labels are not entirely stacked over each other, there is a good chance that they will get imaged. Also, images could be captured by video.

b) Applicability

The imaging station will be mounted with a camera array of a number of webcams, so placing the specimen there will require careful movement. This can probably be achieved by conveyors, which would also facilitate video capture.

c) Expected benefits

Benefits are similar to those of the previous approach. However, in this method, there is a better chance of getting coverage of the labels other than the top label. This approach may also work better for large-winged insects.

d) Difficulties and limitations

There will be a large number of images from varying angles, and their viewing will require time in transcription. This can be improved by image processing that turns the images the right way and corrects the viewing angle.

e) Recommendation

This is a low cost option which will certainly yield valuable experiences, and was tried by LUOMUS during the ICEDIG project, see below.

4.3. Imaging of unit trays

a) Features

This approach is similar to the previous one described in 4.2, but instead of placing individual insects in the imaging station, entire unit trays are imaged. (“Unit trays” are small boxes or trays contained in drawers of collection cabinets, and are being used in most major insect collections. They can be quite different in size and form.)

Individual insects would not be handled. Tagging the individual insects with unique identifier labels would be deferred to a later stage. The labels could be printed on a sheet which would be placed under the unit. The units would be labelled with identifiers as well, which would facilitate rapid retrieval of their data, when the specimens need to be curated.

One or a few shots are made from above and any number of shots from the side using small webcams.

It follows that the images contain many specimens. From the top image the individual specimens can often be automatically picked up using image processing (segmentation). After this step, their

positions in the unit are known, which may assist in automatic segmentation of also the labels from the images made by the webcams.

b) Applicability

This approach is widely applicable for any insect collection which already has been organised in unit trays. It fits well with conveyor-driven imaging.

c) Expected benefits

As this eliminates all handling of individual insects, this approach would achieve the required ten-fold speed increase, and probably more. As such it necessarily is worthwhile to try.

d) Difficulties and limitations

This approach requires heavy computation in the segmentation of the top images, and in a possible creation of a 3D model of the unit, and in extraction of label images of many specimens. The resolution of the top image of each specimen will be lower than those in the previous approaches.

There usually are multiple sizes of unit trays in each drawer. It would be impractical to digitise only some units in a drawer. This can be a problem if conveyors are being used to move the units in the imaging station, and the imaging station cannot handle unit trays of different sizes. Therefore multiple parallel conveyors of various widths may need to be used.

Printing and attaching labels with unique identifiers to all specimens in the unit trays will be challenging, and if postponed to future, will require careful instructions for the curators. This is a major complication in this approach.

Limitations of depth of field and focus point can also become problematic in this approach.

e) Recommendation

This approach seems to offer a large benefit, and should be tried. There are some technical obstacles in the image processing, but these can probably be handled using available technology. This approach was tried by LUOMUS during the ICEDIG project, and the experiences are described below.

4.4. Camera in robot arm

a) Features

The above approaches use fixed cameras, and are suitable for installation in a conveyor setup. A different approach would employ only one camera, which would be installed on a robot arm. The camera would take a large number of shots from different angles of the specimen that would be mounted in a stand. This would be quite similar to the ZooSphere system, but not aim for precise digitisation of insect specimens for 360-degree viewing, and hence only require a few good shots that can be taken fast.

In a different variant, which could be tried after the system works, the robot would work on a unit or on an entire drawer.

The key feature of this approach would be the communication between the robot arm and the camera. Such systems are being used in medical and industry applications. Ideally, the system would need to understand what it sees and steer the imaging in real time.

In an extreme variant the robot arm would not only carry a camera, but an instrument to space out the labels as needed for good imaging.

b) Applicability

This approach might fit all cases of imaging insect collections. However, these cases (individual insects, units, drawers) should probably be treated differently, but at the moment we do not have enough knowledge to specify them in such detail.

c) Expected benefits

This could become one-size-fits-all solution for imaging insect collections. The robot could be left alone to do imaging 24h/365d. Only loading new units and drawers would require small breaks.

There also is potential for automated focus stacking. When the camera is in the right position, instead of just one shot several photographs with varying focus could quickly be made.

d) Difficulties and limitations

Industrial-strength robots are still expensive. There are also cheap robots, but they will not be applicable in these kinds of situations, or need to be amended for the situation

The communication between the robot arm and vision system requires an advanced data processing system. These are probably available from research and industry, but will require adjustment and testing. There may be a high cost in acquiring such a system.

e) Recommendation

This approach should be tested in cooperation with an advanced robotics and machine vision lab.

4.5. Cameras on rails

a) Features

This approach is similar to the previous one, but does not employ a robot arm. Instead the camera is placed on rails moving on one or two axes, and which would work over a drawer. This is basically the SatScan system, but adds the capability to tilt the camera to also see the labels. A smaller, simpler variant would work only on a unit.

b) Applicability

This approach works on any type of insect drawers and units.

c) Expected benefits

The physical setup is not expensive. Therefore, many systems could be installed in parallel to work overnight to produce images of tens of drawers. In that sense, this is an alternative to conveyors. Human effort is probably smaller than operating conveyors.

c) Difficulties and limitations

It is time-consuming to capture large amounts of images during the imaging process. It generates huge numbers of high-resolution images requiring large high-performance storage. It is computationally expensive to process those images.

e) Recommendation

This approach should be tested in cooperation with an advanced robotics and machine vision lab. LUOMUS subcontracted testing of this approach to NampaWorks Ltd, and the results are presented below.

4.6. Terahertz time-gated multispectral imaging

a) Features

Imagine reading a book without opening it, seeing ink through the paper… Terahertz technology has recently been introduced to airport security screening of passengers, and can visualise any objects hidden in pockets and elsewhere. Redo-Sanchez et al. (2015) describe how they extracted occluding textual content from a packed stack of paper pages down to nine pages without human supervision. They achieved this through time-gated terahertz scanning. Their application is close enough to our target application of reading stacked labels from pinned insects, and possibly through the wings of spread specimens.

For a time-gated use, the object that will be studied would need to be installed in a motored environment, so that the layers would be imaged separately. This would probably require placing the pinned insect on a stand, and then moving the stand by a motor at millimetre steps across the range of stacked labels. Alternatively, the scanner could be moved in a similar fashion.

b) Applicability

It is not yet known how the scan would react to insect wings and insect bodies, but labels can possibly be read.

c) Expected benefits

No need to handle the labels.

d) Difficulties and limitations

The resolution of what can be read is related to wavelength, which is about one millimetre. In insect labels the text is very small and may not be readable. Workarounds need to be investigated.

Motorised movement across the layers can take time, as in stack imaging. Again, this needs to be investigated.

e) Recommendation

This is a promising new technology that should be tried with a collaborating organisation that has the required equipment.

5. Results of the experiments

Three of the listed experiments, 4.2, 4.3, and 4.5, described in Section 4 were selected by LUOMUS to test the feasibility and potentials in digitisation of pinned insects. Moreover, the integration into the existing conveyor belt driven pinned insect digitisation system (Tegelberg et al. 2014, 2017), see its setup in Fig. 1 was explored for experiments 4.2 and 4.3.

Figure 1. The conveyor belt driven pinned insect digitisation system (Tegelberg et al. 2014, 2017).

The main goal of both experiments is to capture the label information as much as possible without manual operations on the label, i.e., removing the labels from the pin. A conveyor belt driven system will be utilised to transport the specimen or unit trays to the imaging zone. This will reduce the efforts and time to prepare the specimens before imaging. This is achieved by utilising a camera array with multiple compact high-resolution webcams. However, experiment 4.2 targets the single specimen, while 4.3 targets the unit tray.

The hardware for experiment 4.2 and 4.3 are almost identical, both utilising the camera-array to make a one-time capture of the specimen (4.2) and the unit tray (4.3). Simultaneously all webcams in the camera-array will be triggered to take multiple very high-resolution images. 4K video shooting is also available. Along with conveyor belts, viewing from various angles of the specimen or specimens in the unit and their associated top-most labels will be achieved, without moving the camera-array and the specimen/unit. The integration of one webcam into the conveyor belt driven insect digitisation system is shown in Fig. 2.

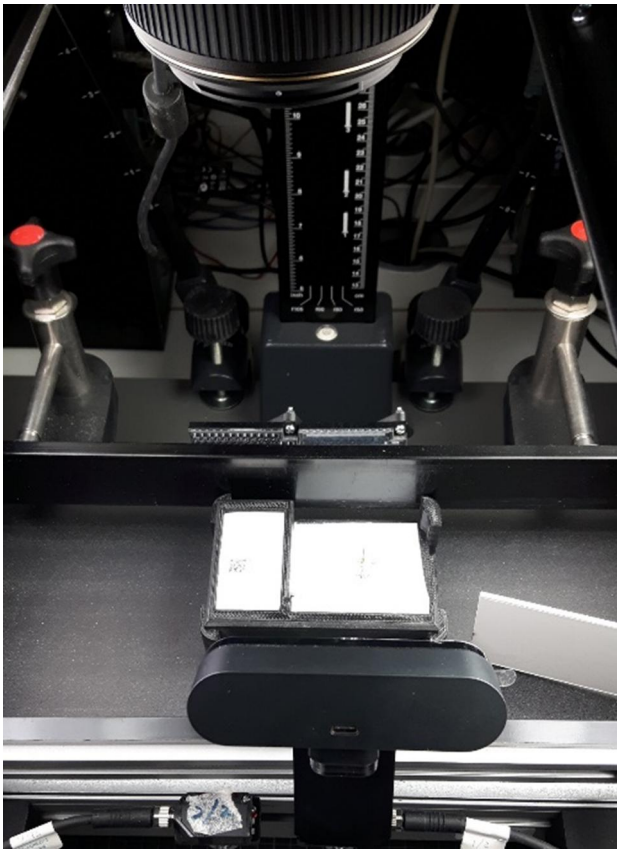
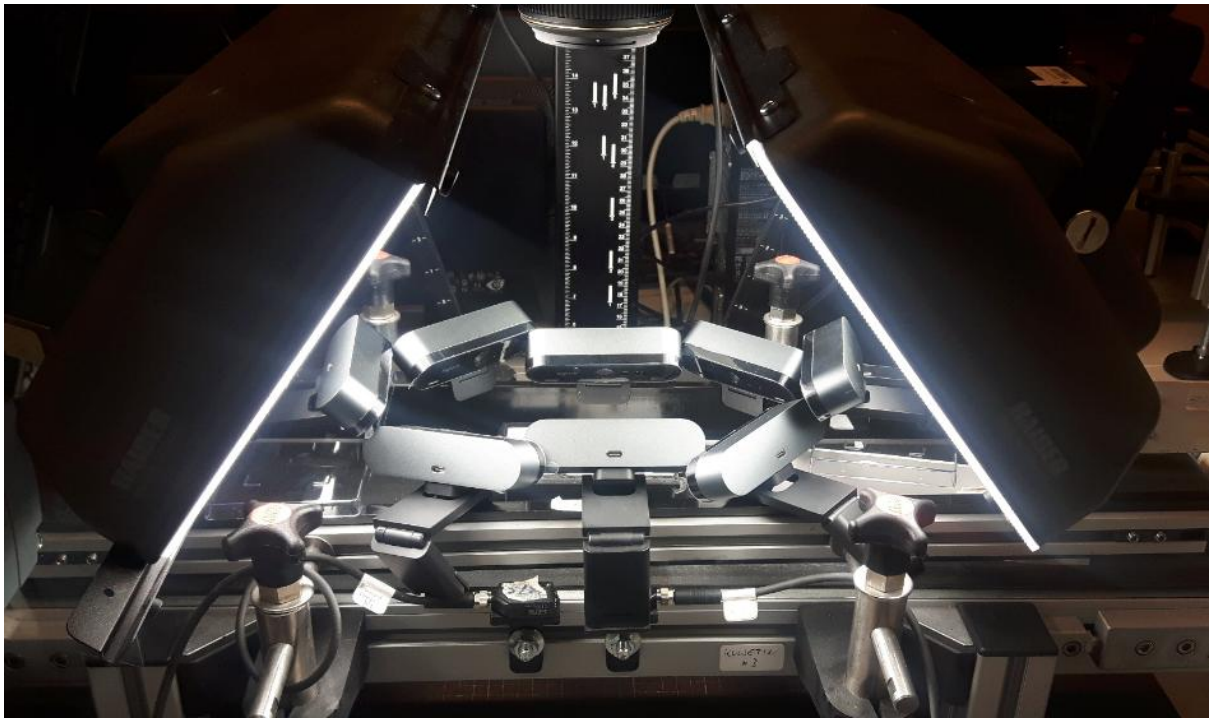


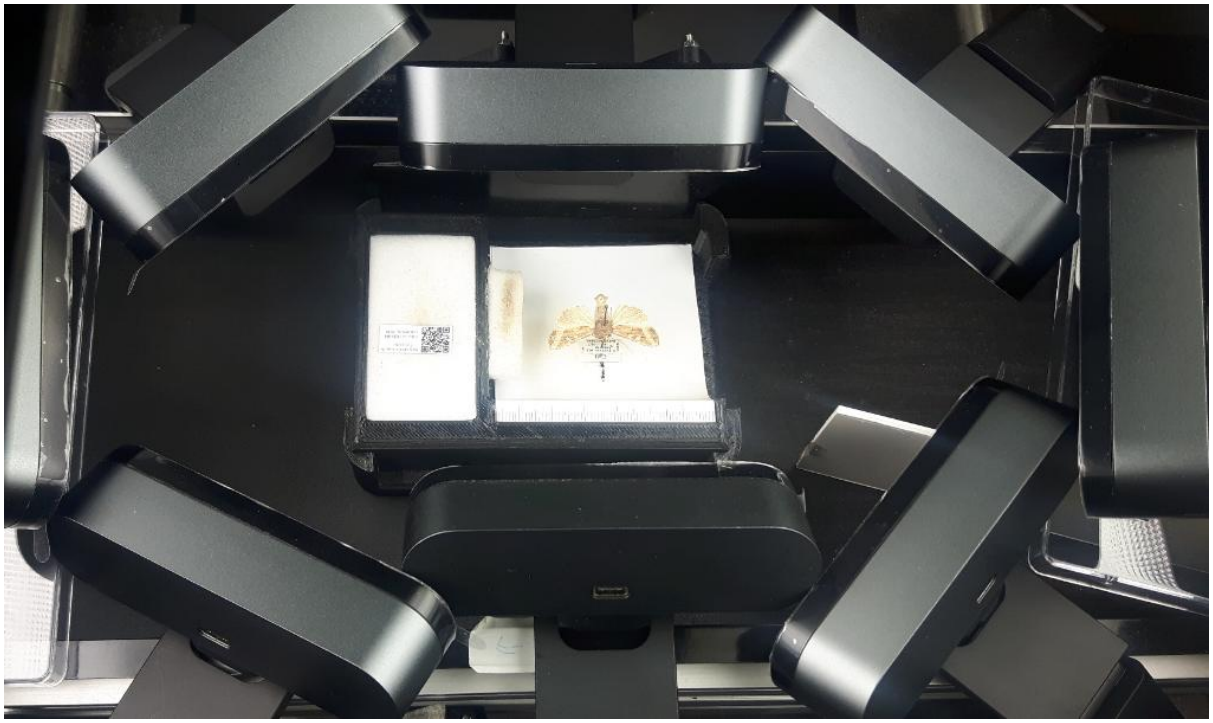
Figure 2. An example of the setup of the imaging area of the conveyor belt driven insect digitisation system with a webcam as the side camera.

Adding layers and webcams in the camera array will increase the chance of capturing more label information. However, costs of the hardware, storage and computation resources will also increase. Also for experiment 4.2 that targets the single specimen, it is easier to record the label information when compared to 4.3 for the whole unit tray, since there is no overlapping of specimens and no occlusion of the label for one single specimen. A single layer camera array with 8 webcams, schematic setup shown in Figure 3, may be capable to handle the single specimen digitisation in experiment 4.2. Figure 4 shows the same setup for the unit tray digitisation in experiment 4.3. It is not necessary to utilise the same settings, the number of layers and webcams, in two experiments. The exact settings may depend on the specimen and the collection type that are to be digitised.

The same camera array setting shown in Figs. 3 and 4 was utilised in both experiments 4.2 and 4.3. The camera array will capture eight side images of the specimen/unit from eight different angles. And the image taken above the specimen/unit will be taken by a single camera.



(a)



(b)

Figure 3. An example of the setup of the imaging area of the conveyor belt driven insect digitisation system with a single layer camera array of 8 webcams for single specimen digitisation show from (a) the side view and (b) the top view.

Figure 4. An example of the setup of the imaging area of the conveyor belt driven insect digitisation system with a single layer camera array of 8 webcams for unit tray digitisation

5.1. Results of experiment 4.2 using multiple webcams

Specimens and labels will be imaged from a number of different angles and directions in this experiment 4.2. The example of images taken from 8 different angles and one from the top is shown in Fig. 5.

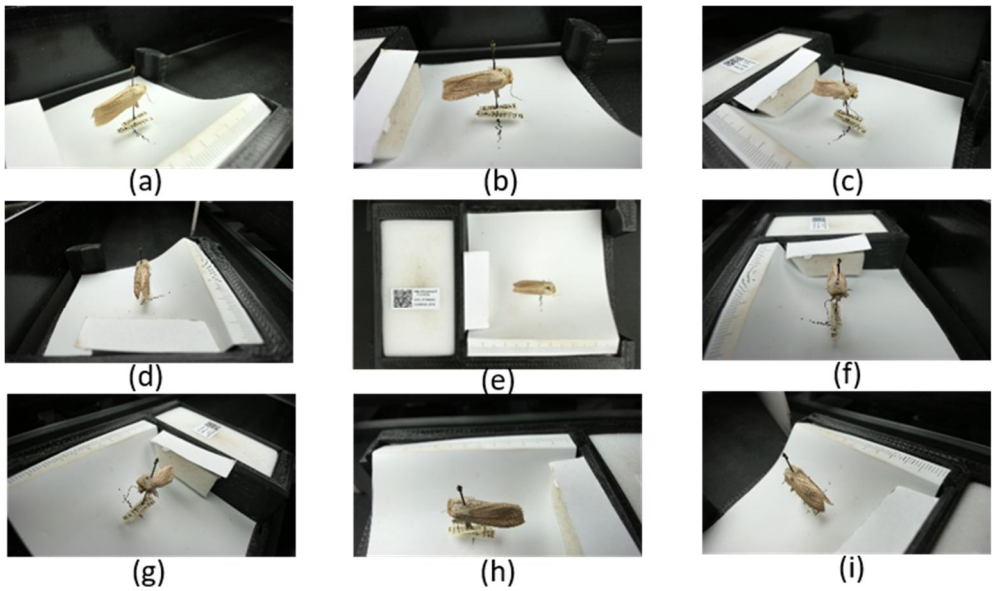


Figure 5. Imaging results of a single specimen, (a)-(d) and (f)-(i) from 8 different angles and (e) from the top.

From Figure 5(e) the top view of the specimen, we can see that the conventional approach, which is imaging of the specimen from above cannot capture the label information beneath the specimen unless the labels are removed from the pin. Therefore, it is necessary to add side cameras as we do in this experiment to capture the label information, but without manually manipulating labels on the pin. Fig, 6 gives an example of the image taken by one of the webcams from the side view, which is a zoom-in version on the specimen from Fig. 5(b). We can see that by using a single side camera the topmost label information can be captured, and other lower-positioned labels as well, if labels are spaced out well.

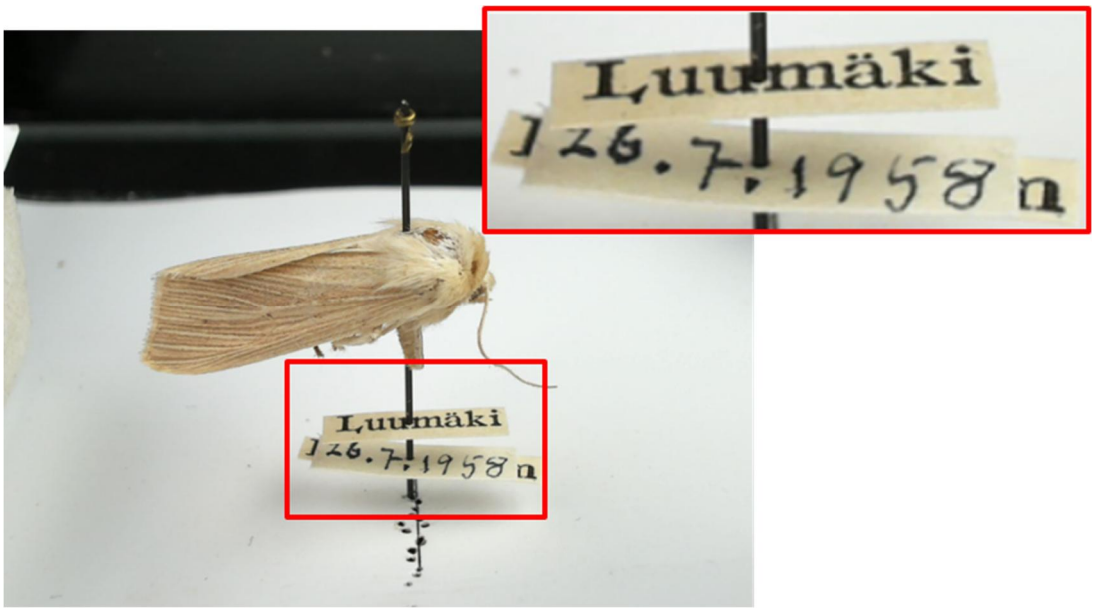


Figure 6. An example of the label area cropped from image of Figure 5(b) taken by one of the webcams from the side view.

Figure 7. An example of the label areas cropped from the two different side view images of (a) Figure 5(a) and (b) 180 degree rotated Figure 5(h).

However, usually there are multiple labels and they are not always spaced out well. From Fig. 6, it can be seen that the second layer label is partial covered by the topmost label and the third layer label is almost totally stacked with the second layer label. By taking images from multiple views, the chance to fully capture the label information will increase, as shown in Fig. 7 where partial covered information on the second label is revealed. But most of the texts on the third layer label still cannot be read out, because it is too close with the second layer label.

Therefore, we tried to do 3D reconstruction from the 8 images taken from the 8 different views of the specimen. We tested Agisoft Metashape[2] photogrammetry software. But the quality of reconstructed 3D model was not good enough, see Fig. 9 for example. Based on the generated 3D model, the labels can be segmented, see Fig. 10. Because not all images are well aligned, the failed aligned images are not included in the 3D reconstruction process. Eight images is a quite low count for photogrammetry, and the resolution of the cameras also plays a role. §By adding more webcams at different positions, the 3D model quality may improve. However, due to the limited space of the image area at the current conveyor belt driven digitisation system, it is not possible to add more webcams. In addition, adding a calibration chart or targets may improve the image alignment to improve the quality of the 3D model.

[2] https://www.agisoft.com

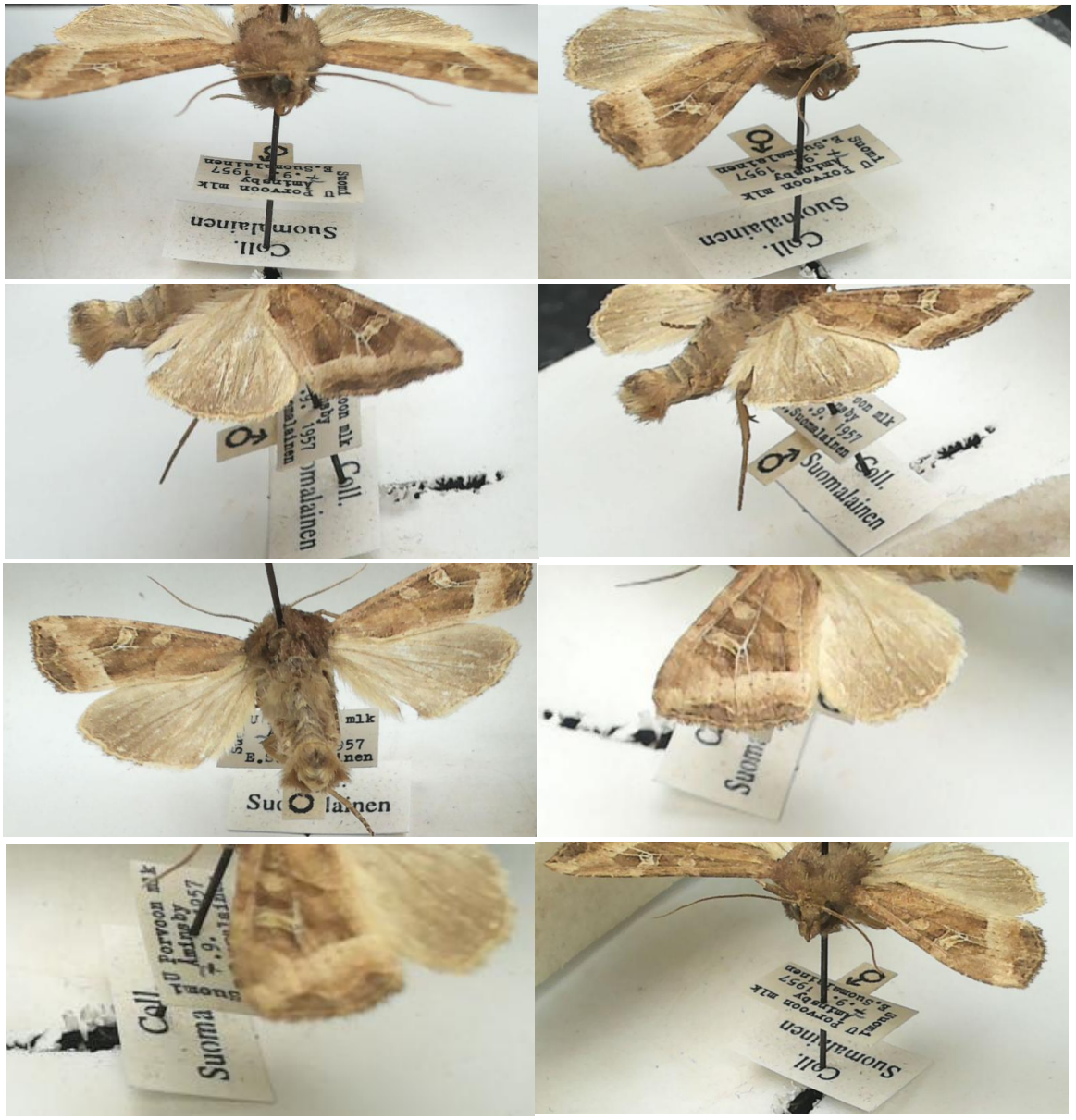


Figure 8. An example of label areas cropped from the image taken by 8 webcams.

Moreover, the label information may need different camera viewers for different types of the specimen and the collection. In addition, the shape and size of the labels, and the alignment of the labels, may influence the visibility of the labels. See example in Fig. 8 that shows the cropped label images from different views from 8 webcams. By increasing the number and layer of webcams, there will be more views of the specimen and the labels from different angels. This will increase the possibility to capture more label information. In addition, this gives the chance to generate 3D model of the specimen and its lower labels. 3D model will not only help to extract label information but also provide more information on the specimen when compared to the convention 2D imaging of the specimen from the top view.

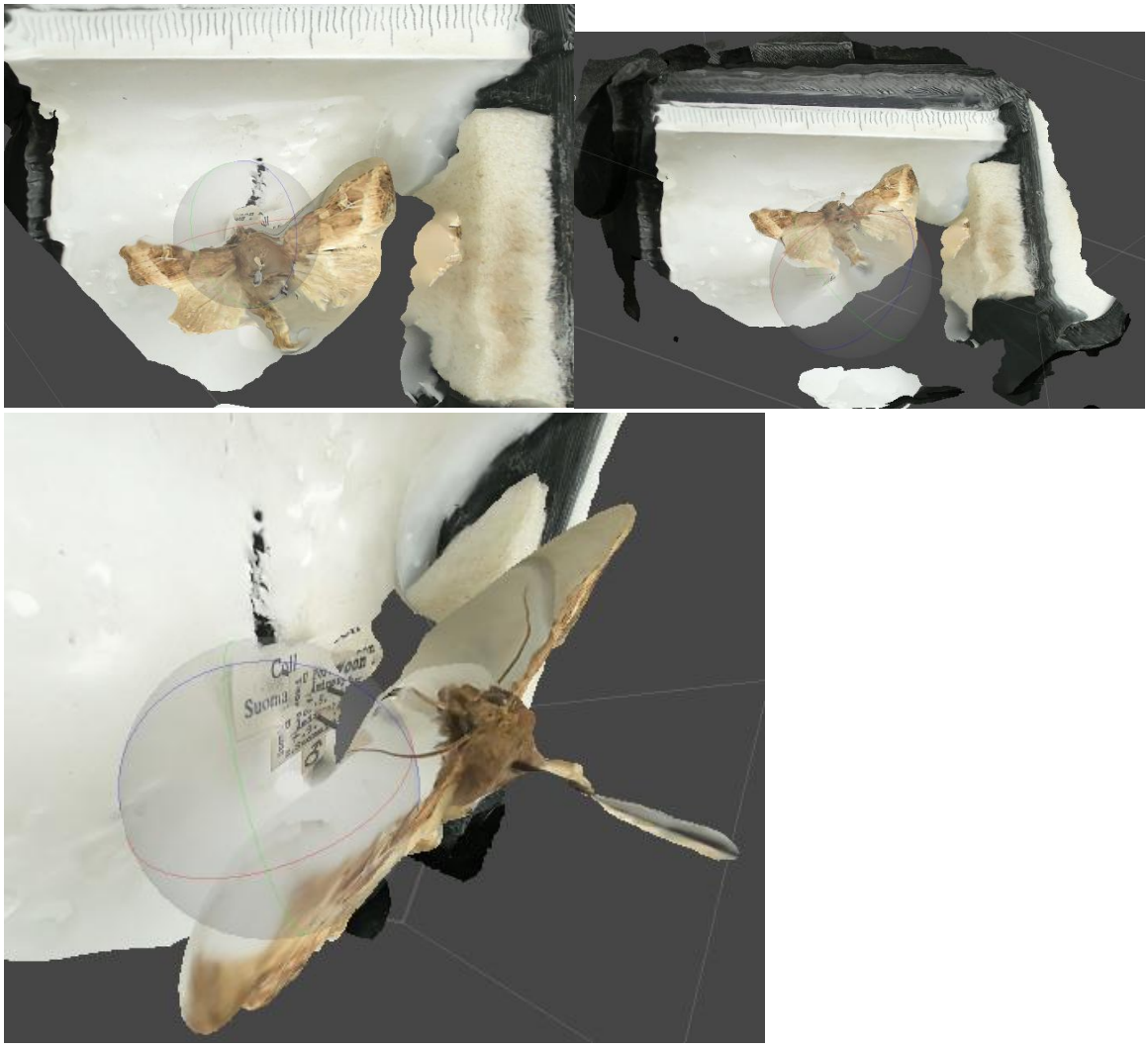


Figure 9. An example of the textured 3D view reconstructed from 8 different images of the specimen and the labels.

From the preliminary experiment results, it is shown that there is potential to use the webcam array to do the digitisation of the small pinned insect specimen and its lower labels along with the conveyor belt system. Usually the topmost label beneath the specimen can be well captured by a single camera from the side view. If the lower labels are spaced out well, the webcam array can capture multiple label information. This will reduce the efforts and time to remove the labels from the pin and therefore to speed up the digitisation process. With the images from multiple views, the label texts can be clearly seen and transcription can be done in the conventional manual way. Moreover, labels can be segmented from the images and Optical Character Recognition (OCR) can be used to extract texts for aiding the manual transcription process or even achieving automated transaction with Natural Language Processing (NLP) techniques. In addition, the quality of the 3D model of the specimen might be achieved by fine tuning the image alignment.

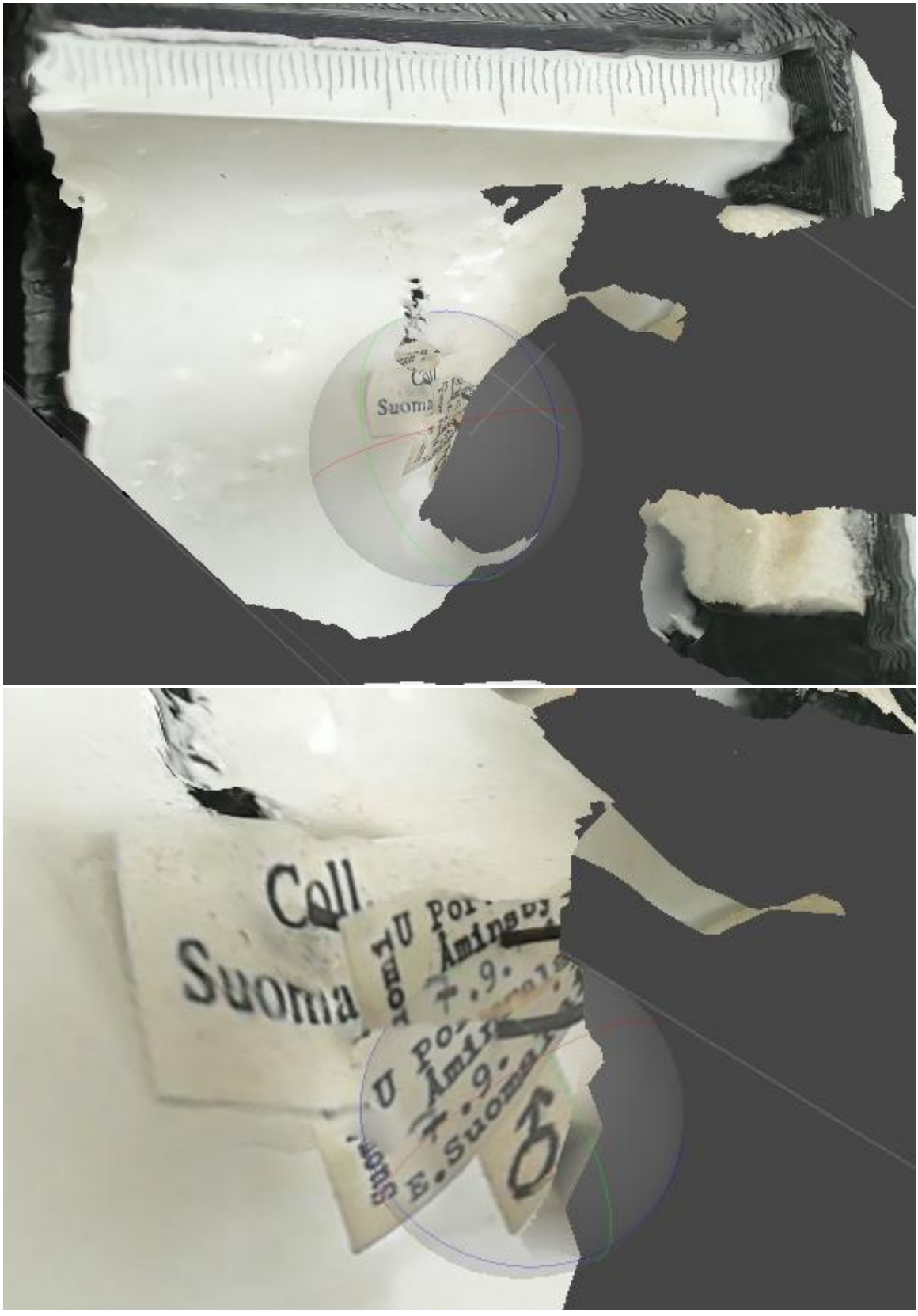


Figure 10. Textured 3D viewing of the labels and its zoom-out view after croping out the specimen. Note the misalignment of images, causing parts of labels to be reconstructed twice.

The estimated performance of the system for the single specimen is up to 360-720 specimens per hour. The technical performance is up to 3600 specimen per hour. However, the bottleneck is how fast the system operator can feed the system. Even when the labels do not need to be removed from the pin in this case, it may still take at least 5-10 seconds in average per specimen for the system operator to pin/unpin the specimen, put the barcode on the tray, and load/unload the specimen tray from the conveyor belts.

5.2. Results of experiment 4.3 imaging of unit trays

Extending the integrated camera array and conveyor belt driven digitisation system from the single specimen imaging to the unit level is not that straightforward. Firstly, in order to use the existing conveyor belts driven digitisation system, the unit size should fit the conveyor belt. Secondly, more webcams are needed in order to capture the full label information because of the possible occlusion of the labels due to the density of the specimens in the unit. Thirdly, the location of the webcams has to be optimised for the dimension of the unit. The preliminary example of images of the unit taken by the same imaging system of 8 webcams from the side views and one from the top is shown in Fig. 11.

From the experimental results, we find that it is more challenging to image the unit when compared to the single specimen. The specimens are arbitrarily located in the unit tray, which makes it difficult to make all objects in focus. Fig. 12 shows one image taken with automatic focus, where the objects at the left bottom corner with red colour bounding box are not in focus, while the objects in the blue colour bounding box is focus. Moreover, even the specimens in the example shown in Fig. 11 are quite sparse in the unit, the labels of the specimen near the corner and border cannot be fully captured. If the specimens are densely located in the unit tray, it will be more difficult to capture the labels beneath the specimens.

We tested the 3D reconstruction from the 9 images shown in Fig. 11 with the same approach in experiment on the single specimen. The quality of the 3D model is also not satisfying, see Fig. 13. This is due to only 5 in 9 images are aligned, resulting in the missing of information from the failed images. From Fig. 14, we can see that the labels in the middle of the unit tray are easy to read. However, for the labels at the corners and near the border walls, the text on the label is difficult to recognise. It may need more captures for the specimen at the corners and near the border walls. Due to the space limitation at the imaging zone at the conveyor belts, it seems to be difficult to add more webcams. It is worth trying to shoot videos of the unit tray while it moving on the conveyor belts. This will increase the chance to capture more views of the specimen in the unit. Accordingly, this will increase the data volume and the computation load on the post processing of the videos.

The preliminary results of the images on the unit tray and the reconstructed 3D models show that if the specimens are not very densely spaced in the unit tray, it seems to be possible to use multiple webcams to capture the label information. A 3D model of the specimens in the centre area of the unit tray can be reconstructed. If the number of images at different views for the specimens at the corner or near the border wall is relatively high, it may also be possible to achieve the 3D model for those specimens.

For the performance of the system on the unit tray, it can be estimated that the number of specimens digitised per hour will be higher when compared with that of single specimens, since

there is no need to pin/unpin each specimen from the unit tray. This will tackle the bottleneck of the digitisation process, the preparation of the specimen before imaging process. Therefore, further investigation on the imaging of units is necessary. Also the barcoding process for the unit level imaging has to be studied.

Figure 11. Imaging results of the unit from the above and 8 different oblique angles.

Figure 12. An image example of camera focus problem.

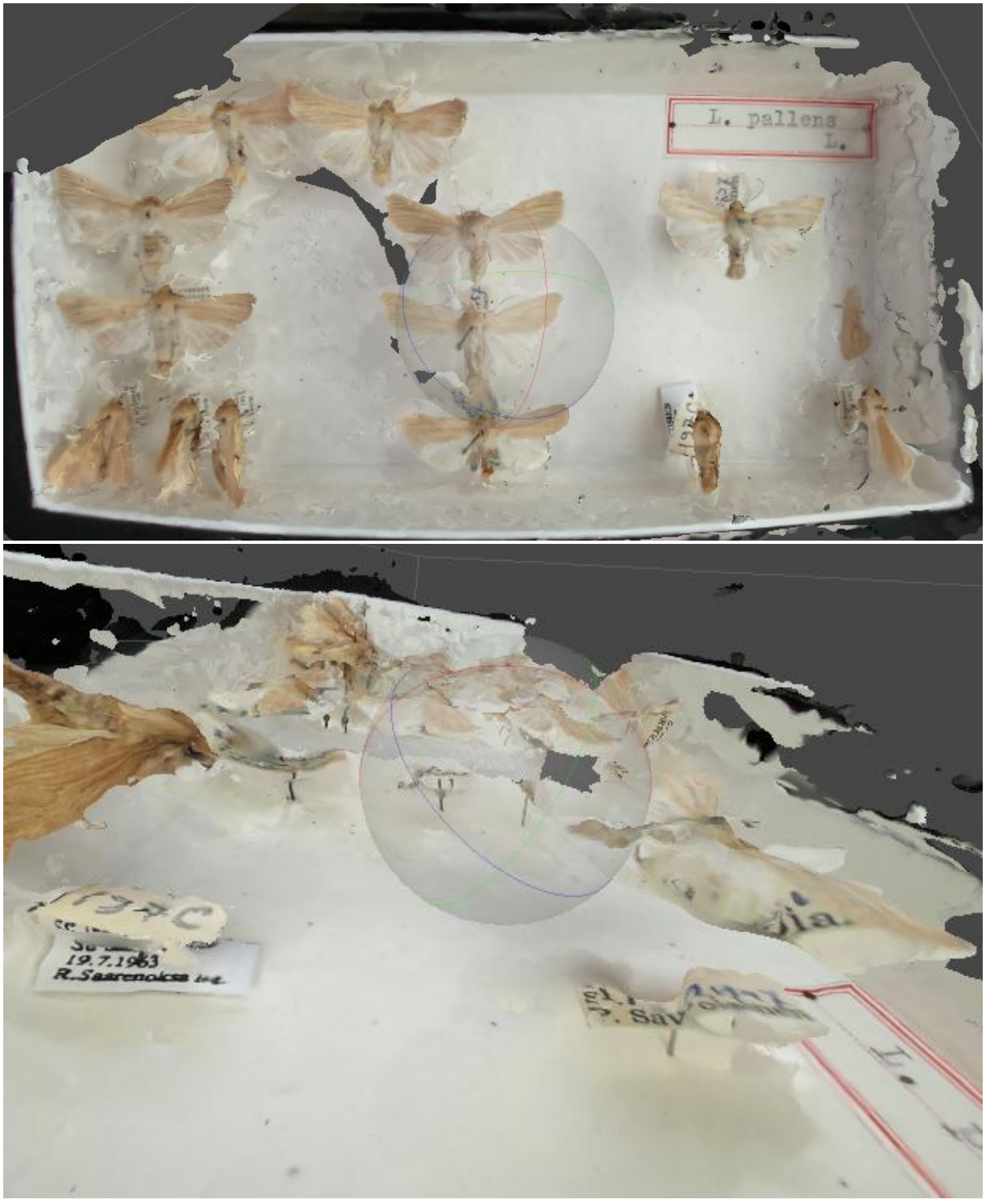


Figure 13. An example of the textured 3D view reconstructed from images shown in Fig. 11.

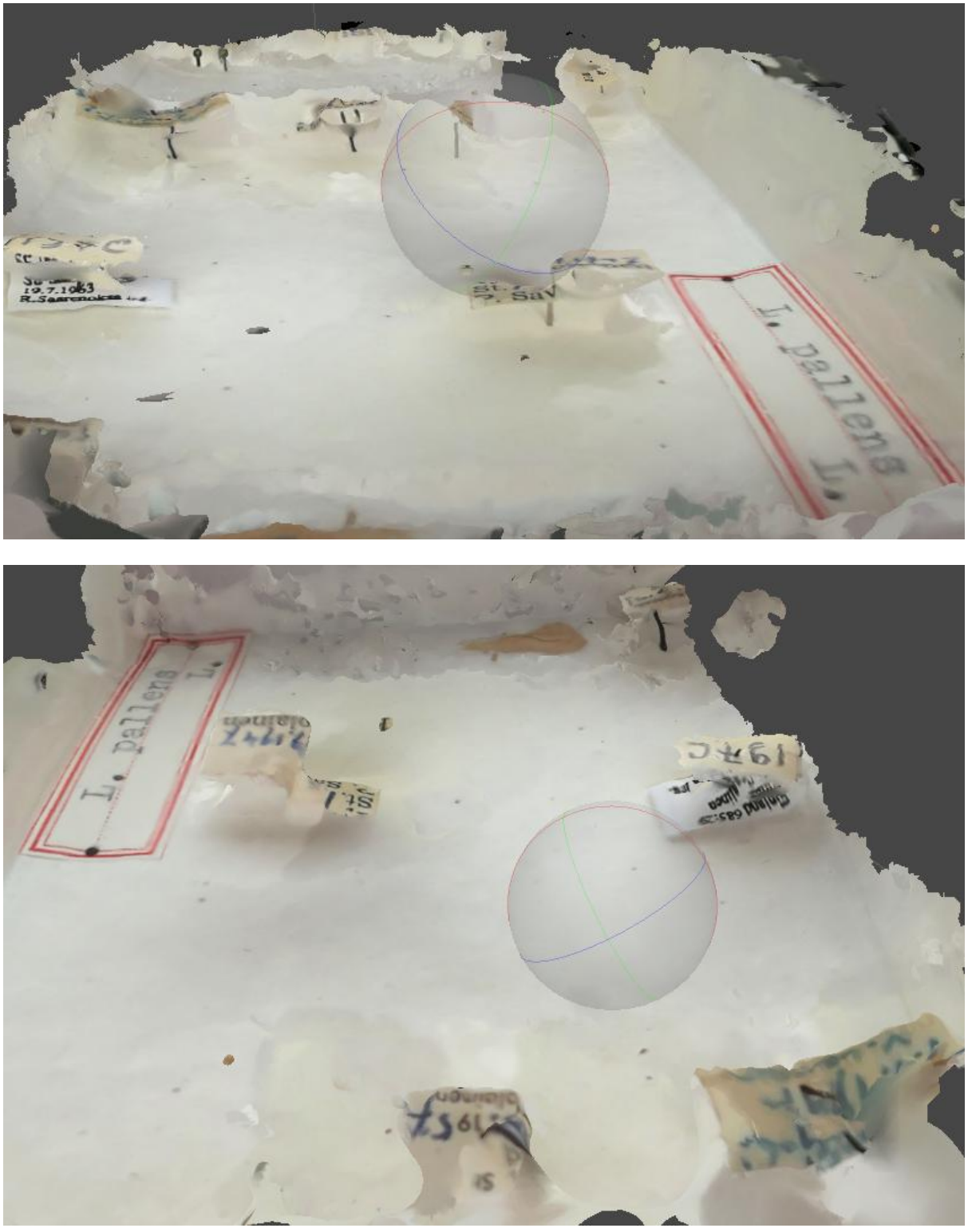


Figure 14. Examples of the textured 3D view at two directions after cropping out the specimens from 3D reconstructed from images shown in Fig. 11.

5.3. Results of experiment 4.5 using cameras on rails

This experiment was carried out by NampaWorks Ltd in consultation and with some guidance by LUOMUS. The system here is called ENTODIG-3D (Ylinampa and Saarenmaa 2019). The imaging station setup is shown in Fig. 15. There is a frame made of aluminium profiles which forms a cube of about 80 cm on each side. On top of this frame, three motorised rails have been mounted, which allows movement across the entire frame in 2 dimensions. In one of the rails a motorised tripod head has been mounted upside down. The tripod head can turn 360 degrees and also tilt up and down360 degrees. In this tripod head, two 4K webcams have been mounted and those are connected to an imaging computer.

A master Python program controls the workflow. In the imaging process, first a 2D-mapping is performed to locate all the specimens. A motorized camera, which is tilted downwards, moves freely over a drawer, taking images. The four corners of the drawer are indicated with QR-codes. Images are stitched together and using TensorFlow object recognition software, as well the insects in the unit tray as their coordinates are then identified and determined (Fig. 16).

Imaging of individual insects is done by moving the webcams above the insect and spinning horizontally 360 degrees above the insect at about 30-degree angle vertically. Because the webcams are somewhat elongated from the axis of the tripod head, this gives a view to the labels under the insect. A total of 30 shots are made, each time rotating the webcams 15 degrees horizontally. One shot is actually a one-second video clip with focus shift.

The video clips are then disassembled into 30 individual frames. These frames are processed with HeliconSoft focus stacking to produce sharp images.

In Agisoft, the mesh (polygon) model calculation is based on the point cloud. When there are fewer points in the point cloud, it is possible to calculate more accurate meshes and textures. On the other hand, textured meshes are easier to detect with TensorFlow object recognition software. Therefore, in order to create more accurate 3D-scans, TensorFlow guides the cropping of a point cloud. Thus, a loop between point cloud, textured mesh and TensorFlow, was created.

All stacked images related to one specimen are fed to Agisoft photogrammetry software, and an overall 3D-model of each insect is created (Fig. 17). All of these specimen 3D-files should have the same scale and orientation.

All of the 3D-files are imported to Blender 3D viewing application. Screenshots (images) are taken from top-view of the 3D-model.

The Python master program takes the files, point cloud and textured mesh, from previous steps and stores those in a temporary database. Because there are multiple specimens in the 3D-model, the “center specimen” must be identified. This is done by feeding the top-view images into TensorFlow, which identifies the detection box in the center of the screen, based on the x- and y-values. The coordinates of “center specimen coordinates” are entered in Blender, that contains all 3D-models of a specimen. With mouse (bot) the center specimen is selected and all other points of the point cloud is deleted. This point cloud is then exported into a new folder called “Cropped specimen.”

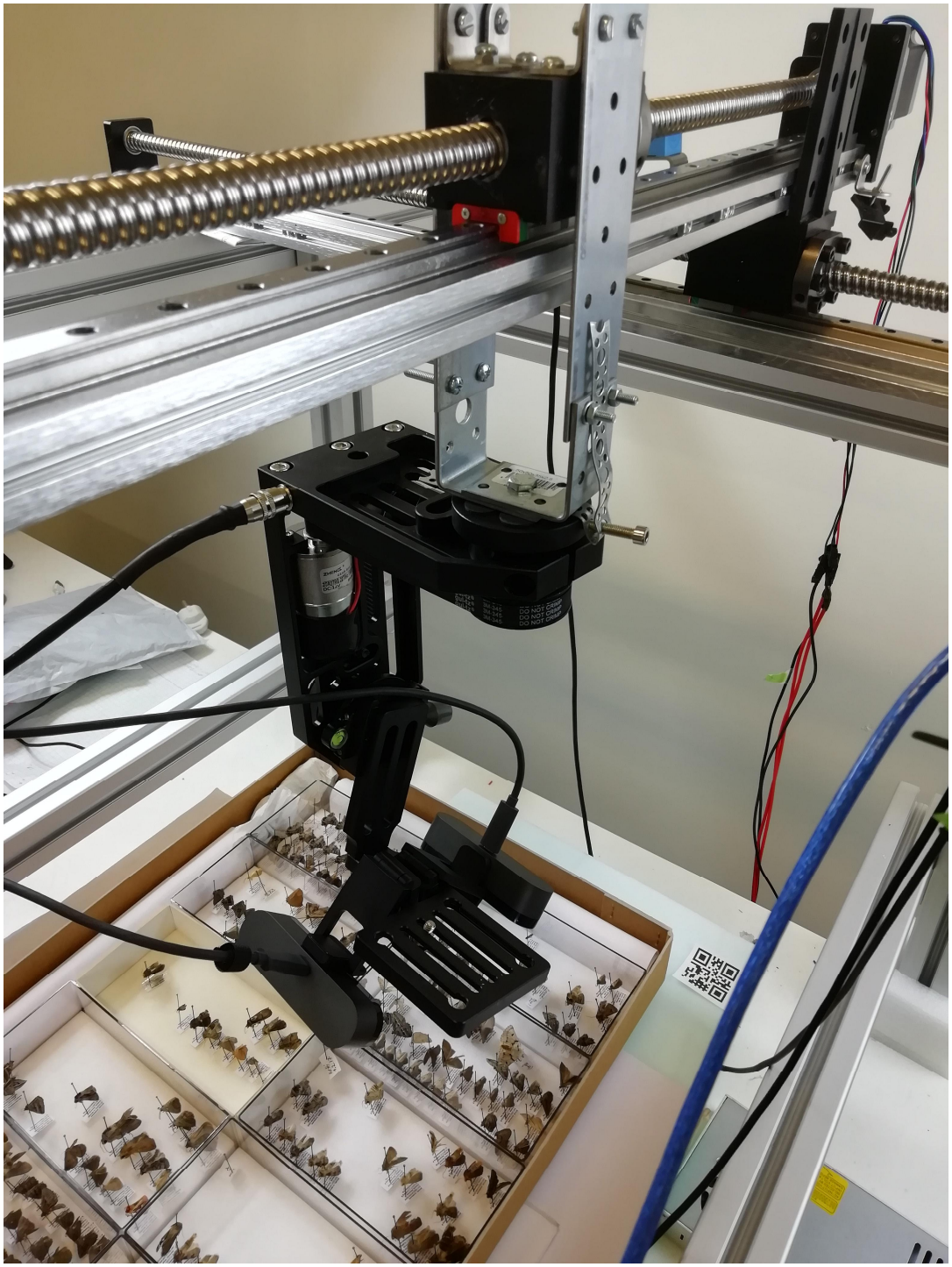


Figure 15. The rail based imaging station.

Now there is only one specimen in the 3D-data (point cloud and mesh), and the label underneath the specimen should be found. As mentioned in earlier phase, an image is taken from the side-view. This is done by importing 3D-data into Blender, and moving the camera 90 degrees to get a side-view in textured mesh-level. Again, TensorFlow finds out the coordinates of the label and the specimen, removing the latter in point cloud –level. To get more detailed mesh, this “vertical crop” point cloud is imported into Agisoft, which calculates a new textured mesh. It is important to set the

texture settings blend to average, because there used to be a specimen over the label. With average, the program calculates the mean value for the texture, leaving the actual label text visible. Finally, an image of the textured mesh of the label can be taken in Blender. Result is an unobstructed view 2D view to the labels and a 3D-model of the insect (Fig. 18).

At this writing, the whole process is still not fully integrated, but all the individual steps have been tested in manual mode. Processing one insect this way takes about 5 minutes (corresponding to 288 specimens in 24 hours). The time goes to 1 minute of making the video clips and 4 minutes back-end processing in creating the 3D models. These bottlenecks could be removed by using more webcams on the tripod head, and using a more powerful compute engine. We estimate that with such improvements the speed of the system could be up to 3,000 specimens in 24 hours. No human operators will be needed, if the imaging system is placed on a large conveyor which feeds the drawers to the imaging.

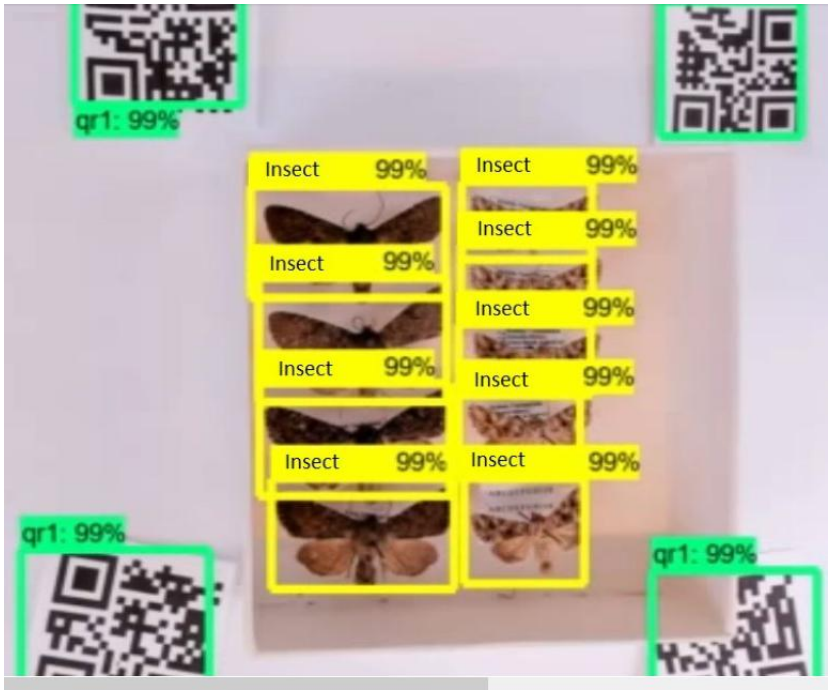


Figure 16. TensorFlow object recognition.

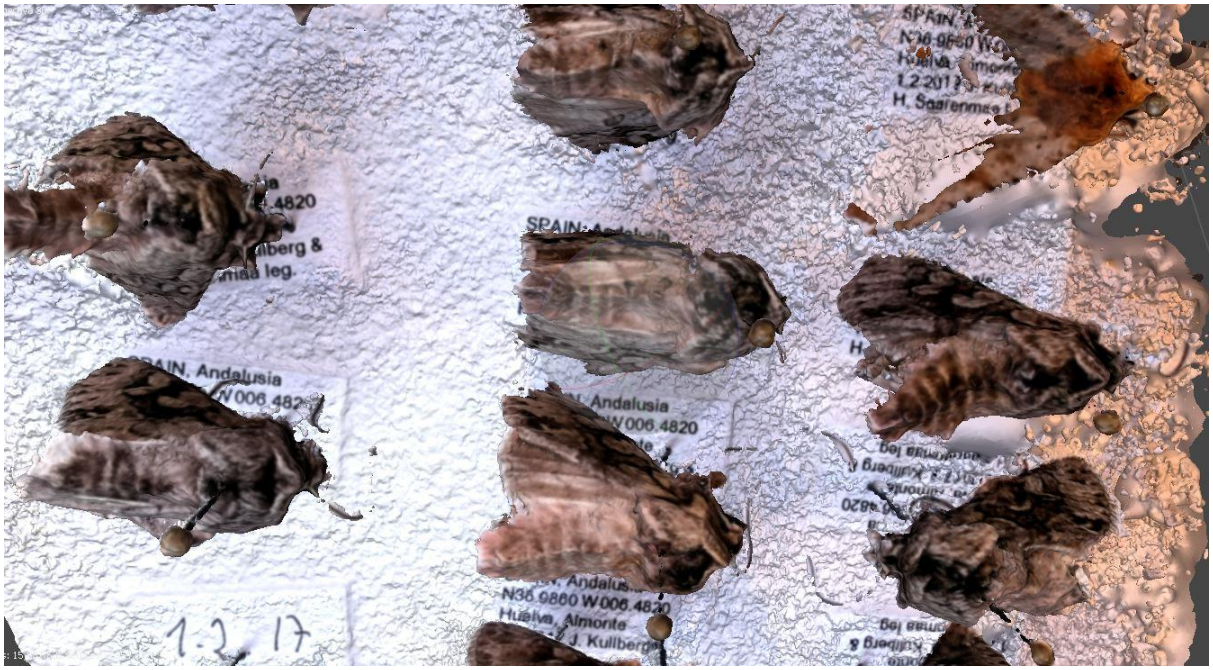


Figure 17. 3D model of a unit tray.

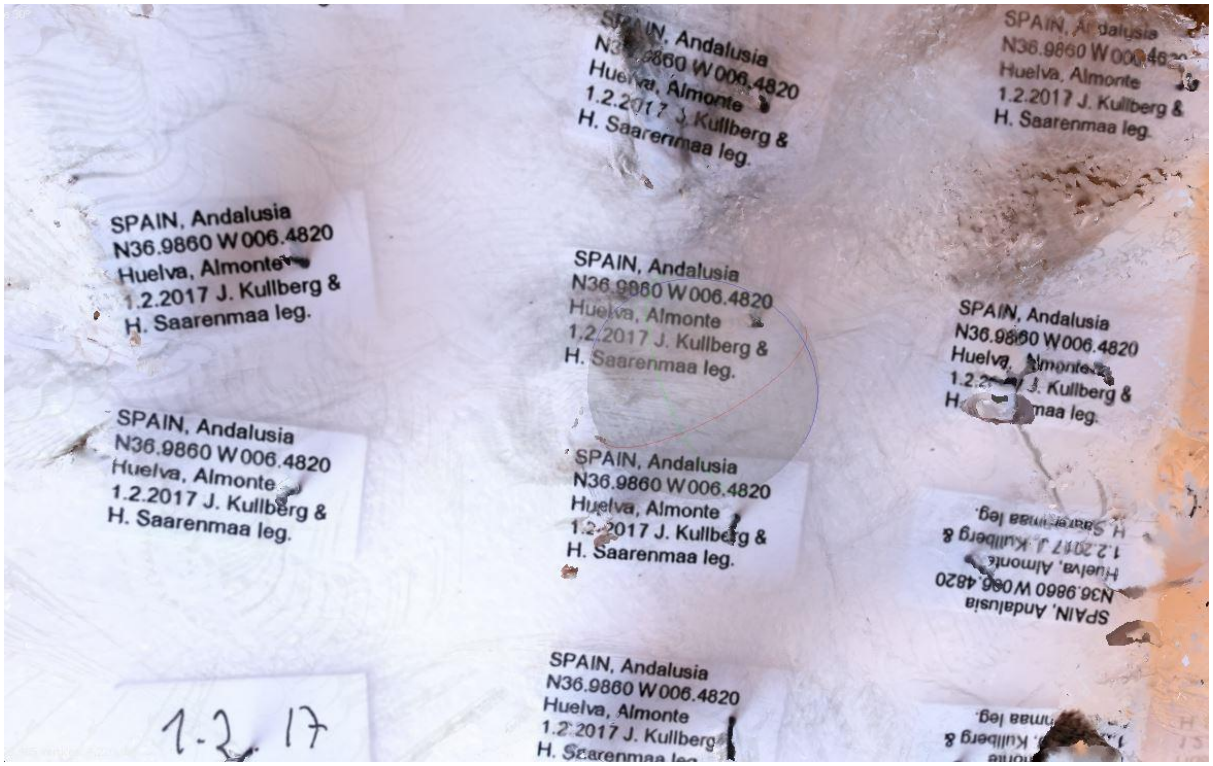


Figure 18. Unobstructed view to the labels in the unit tray shown in Fig. 17. This was achieved by vertically cropping away the insects.

6. Discussion

We still do not have a definitive answer how we could boost the pace of digitisation of pinned insects by a factor of ten. Nevertheless, there are promising developments that may yield the solution in near future.

3D modelling is increasingly being used to digitise cultural heritage (e.g., Ritz et al. 2018). However, none of these solutions have been designed with mass-production in mind. To modify them for mass production might be possible, but will require redesign. Starting from a functioning system and redesigning it is often more efficient than starting from scratch.

Ultimately, it is question of cost. We need to reduce the number of human operators as much as possible. A fully automated line where humans only need to bring the insect drawers would be ideal. We can already see how that might work in the ENTODIG-3D solution mounted on a conveyor.

Another approach to reduce the costs would be a really inexpensive, but fully automated imaging station. If the cost is well below 10,000 €, we could install dozens of those in parallel in each museum. Human operators would only bring one drawer to imaging once or twice a day. Even the biggest collections only have about 50,000 drawers, containing up to a million unit trays. Ten imaging stations operating in parallel 24h / day for twenty years would do the job.

In these fully automated scenarios one problem remains: How to attach unique ID such as QR-code to each specimen? It does not take much time to attach them, but still would mean that each insect is handled by a human for a few seconds. Just attaching codes is much simpler and faster than handling labels off and on, and a typical drawer of 200 specimens can be equipped with QR-codes in about one hour. Deferring this operation to later time bears the risk of error and requires training of all curators and visitors who might handle the drawer or unit tray, and will probably fail. QR-codes need to be about 8-10 mm and need to stick out of the specimen to be readable, and will take up space in the unit tray , which may already be packed too tight to allow this. They also increase the danger of damaging nearby specimens when handled. Therefore many curators shun the use of printed QR-codes. One possibility that could be explored is to make the image of the insect itself the unique identifier! An image taken from a standardised position could be hashed into a bitstream and a have shortcut (database key), which would be the identifier. If we look close, and like human faces, not any two insect specimens are identical. Great progress has already been made in automatically identifying insect species from each other automatically (Valan et al. 2019), so what about individuals? We recommend testing this approach in pilot projects.

The big question is that do we really want to make pictures of 1 billion insect specimens? Would it be better to just transcribe the labels? This is actually what most curators prefer, because in many insect groups a picture of the specimen is not sufficient to determine the species, but examination of its genitalia is required. We are not trying to answer this question is this report, but just point out that transcription of the labels, if done without imaging, must necessarily be done in-house at the institution which owns the collection. This defies the possibility to do the transcription in other countries, where labour costs could be lower and local knowledge of language, geography, and taxonomy often is. So, indeed, we must make pictures of all the specimens, in particular of their labels!

7. References

Adcock M, Nguyen C, Lovell D, La Salle J (2014) Accelerating Entomology with Web3D Insects. 1 p. Web3D 2014 - Nineteenth International ACM Conference on 3D Web Technologies, August 08-10, 2014, Vancouver, British Columbia. DOI: 10.1145/2628588.2635851. https://dl.acm.org/citation.cfm?id=2635851

Agarwal, N., Ferrier, N., & Hereld, M. (2018, March). Towards Automated Transcription of Label Text from Pinned Insect Collections. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 189-198).

Blagoderov V, Smith V (2012) No specimen left behind: mass digitization of natural history collections. ZooKeys 209. https://zookeys.pensoft.net/browse_journal_issue_documents.php?issue_id=361

Brecko J , Mathys A (2016). Handbook of best practice and standards for 3D imaging of natural history specimens. SYNTHESIS3 Project Deliverable. Retrieved April 10, 2018: http://biowikifarm.net/v-mfn/3d-handbook/3d_Imaging_Handbook:Main_Page

Häuser, C.L., Steiner, A., Holstein, J. & Scoble, M.J. (eds.) (2005): Digital Imaging of Biological Type Specimens. A Manual of Best Practice. Results from a study of the European Network for Biodiversity Information. Stuttgart. viii + 309 pp. ISBN: 3-00-017240-8. https://www.gbif.org/document/80576/digital-imaging-of-biological-type-specimens-a-manual-of-best-practice

Heerlien M, van Leusen J, Schnörr S, de Jong-Kole S, Raes R, van Hulsen K (2015) The natural history production line: An industrial approach to the digitization of scientific collections. ACM J. Comput. Cult. Herit. 8, 1, Article 3 (February 2015), 11 pages. DOI: http://dx.doi.org/10.1145/2644822

Hereld M, Ferrier NJ, Agarwal N, Sierwald P (2017) Designing a high-throughput pipeline for digitizing pinned insects. 2017 IEEE 13th International Conference on e-Science (e-Science), Auckland, 2017, pp. 542-550. doi: 10.1109/eScience.2017. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8109193&isnumber=8109104

Hereld M (2018) BugSnapper: Camera Design for High-Throughput Digitization. Email Correspondences

Holovachov O, Zatushevsky A, Shydlovsky I (2014) Whole-drawer imaging of entomological collections: Benefits, limitations and alternative applications. Journal of Conservation and Museum Studies. 12(1). http://doi.org/10.5334/jcms.1021218

Hudson, L. N., Blagoderov, V., Heaton, A., Holtzhausen, P., Livermore, L., Price, B. W., ... & Smith, V. S. (2015). Inselect: automating the digitization of natural history collections. *PLoS one*, *10*(11), e0143402.

Kroupa A, Pluta M, Schurian B, Glöcker F (2014) ZooSphere - A tool for automated spheric image capturing and interactive 3D visualization of biological collection objects. The Society for the Preservation of Natural History Collections annual conference, June 22-27, 2014, Cardiff, UK. https://www.researchgate.net/publication/268779382_ZooSphere_-_A_tool_for_automated_spheric_image_capturing_and_interactive_3D_visualization_of_biological_collection_objects

Kutner M (2014) Mission Not Impossible: Photographing 45,000 Bumblebees in 40 Days. https://www.smithsonianmag.com/smithsonian-institution/mission-not-impossible-photograph-thousands-bumblebees-40-days-180952211/#xJBZv7VJ3OoCw1C8.99

Nieva A, Rosin P, Sun X (2018) Specifications of tests on automated methods for mass 3D imaging. H2020-INFRADEV-2017-1 Project ICEDIG, Grant Agreement Number 777483. Milestone MS16 report, 11 p. Cardiff University.

Oever JP, Gofferje M (2014) From pilot to production: Large scale digitization project at Naturalis Biodiversity center. ZooKeys 209, 87-92. https://zookeys.pensoft.net/article/2921/

Price, B. W., Dupont, S., Allan, E. L., Blagoderov, V., Butcher, A. J., Durrant, J., ... & Smith, V. (2018). ALICE: Angled Label Image Capture and Extraction for high throughput insect specimen digitisation. https://doi.org/10.31219/osf.io/s2p73

Redo-Sanchez A, Heshmat B, Aghasi A, Naqvi S, Zhang M, Romberg J, Raskar R (2015) Terahertz time-gated spectral imaging for content extraction through layered structures. Nature Communications 7, 12665. http://www.nature.com/articles/ncomms12665

Ritz M, Knuth M, Santos P, Fellner DW (2018) CultArc3D_mini: Fully Automatic Zero-Button 3D Replicator. 10 pages. In: R. Sablatnig and M. Wimmer (Editors), EUROGRAPHICS Workshop on Graphics and Cultural Heritage. DOI: 10.2312/gch.20181335.

Schmidt S, Balke M, Lafogler S (2012) DScan – a high-performance digital scanning system for entomological collections. ZooKeys 209, pp. 183-191. https://zookeys.pensoft.net/article/2913/

Ströbel B, Schmelzle S, Blüthgen N, Heethoff M (2018) An automated device for the digitization and 3D modelling of insects, combining extended-depth-of-field and all-side multi-view imaging. ZooKeys 759: 1-27. doi: 10.3897/zookeys.759.24584

Tegelberg R, Mononen T, Saarenmaa H 2014. High-performance digitization of natural history collections: Automated imaging lines for herbarium and insect specimens. Taxon 63 (6) 1307-1313. http://www.ingentaconnect.com/content/iapt/tax/2014/00000063/00000006/art00010

Tegelberg R, Kahanpää J, Karppinen J, Mononen T, Wu Z, Saarenmaa H 2017. Mass digitization of individual pinned insects using conveyor-driven imaging. In: Hereld M (Editor) High throughput digitization for natural history collections. 2017 IEEE 13th International Conference on e-Science (e-Science 2017). Auckland, New Zealand, 24-27 October 2017. 5 p. https://ieeexplore.ieee.org/document/8109190/

Valan M, Makonyi K, Maki A, Vondráček D, Ronquist F (2019) Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. Systematic Biology, syz014, https://doi.org/10.1093/sysbio/syz014

Ylinampa T, Saarenmaa H (2019) Photogrammetrical insect collection capturing system (ENTODIG-3D). Biodiversity_next conference, Leiden, the Netherlands 2019-10-21/25. Abstract (in print).