Designing a high-throughput pipeline for digitizing pinned insects

Mark Hereld[1,2], Nicola J. Ferrier[1,2,3], Nitin Agarwal[4] and Petra Sierwald[5]
[1]Mathematics and Computer Science Division
[2]Computation Institute, University of Chicago
[3]Institute for Molecular Engineering, University of Chicago
[4]University of California, Irvine
[5]Field Museum of Natural History
Argonne National Laboratory, 9700 S. Cass Ave., Lemont, IL, 60439, USA
{hereld, nferrier} at anl.gov, agarwal at uci.edu and psierwald at fieldmuseum.org

***Abstract*—This paper presents the design and prototyping of hardware and software to address the problem of rapid and reliable 3D digitization of very large collections of pinned insects. Using the collection at the Field Museum of Natural History (FMNH) as a use case, a pipeline to ingest the entire collection of 4.5 million specimens in circa 1-2 years imposes a few second limit on average processing time per specimen. We describe the design and implementation of multi-camera systems capable of rapidly capturing light field imagery for 3D reconstruction of label surfaces and specimen in single snapshots consistent with this time constraint. With imagery captured using these prototype multi-cameras we demonstrate methods under development for 3D reconstruction of pinned insect specimens and for processing text on label surfaces.**

I. INTRODUCTION

Large-scale collections of objects, which can comprise millions to hundreds of millions of specimens, provide data for studies of taxonomy, biodiversity, invasive species, biological conservation, land management, pollination, and biotic responses to climate change. These collections represent a significant societal investment in research and applied environmental science. However, they also present extraordinary challenges to researchers, archivists, and curators.

Working with the FMNH, we are developing a means to rapidly digitize such large collections. Transporting objects between the digital and physical realms is an increasingly common activity in the worlds of science, industry, and entertainment. Consequently a great deal of technology has been developed in support of these processes. In the scientific arena, digital information and visualization of physical objects often creates new opportunities for discovery and understanding. Once in the digital realm, objects are impervious to time, available for richly informative and powerful query-based exploration and analysis methods, accessible to a much wider community of researchers, and deployed in a wide range of outreach activities. In aid of this goal, we are developing a multi-camera head designed for single-shot digital capture of casually aligned pinned insects. Complementing this hardware design and prototyping effort, we are developing methods for label capture for input to optical character recognition (OCR) and three-dimensional (3D) model capture in the face of uncertain orientation, position, and occlusion. Digital capture of 3D models can provide opportunities for research and education without the need to handle specimens and camera based approaches may yield cost-effective, natural color models [10].

In the sections that follow, we will describe our approach to addressing the challenges of this problem. To guide that discussion, it is important to list the priorities as we see them. Our first level concerns are to *digitally tag* each specimen, *capture images of all of the labeling* suitable for digitizing by automatic or other means, and capture *reference imagery* suitable for identity verification. Our next most important concern is to enable ingest of the entire collection of pinned insects in one to two years, short enough to demonstrate success in the course of a typical federally funded project. Finally, because of the potential value of detailed digital data for scientific discovery, we aim to capture imagery amenable to detailed study, providing a *quantitative virtual stand-in of the actual object* for species identification and photogrammetry. This will include data capture for generating 3D models of each specimen suitable for scientific study and outreach activities.

II. DESIGN CONSIDERATIONS

The problem of digitizing such a complex and diverse collection of objects seems to require a flexible and multi-faceted solution.

A. Challenges of high throughput digitization

We discuss the problem of developing a system and process capable of digitizing a large pinned insect collection in terms of five concerns: estimating the required speed for a problem of this scale and its immediate consequences; identifying characteristics of the objects that will determine how much time it will take to handle them; reliably imaging the label surfaces for casually aligned specimens; determining what needs to be computed in real time and what can wait for offline analysis; and, designing a flexible integrated pipeline that will address these issues. We discuss each in turn in the paragraphs to follow.

High throughput. The FMNH, for example, has a pinned insect collection with an estimated 4.5 million specimens in approximately 15 thousand drawers. In addition to being compatible with project funding practicalities, the goal of digitizing the collection over the course of a year or two has additional benefits. First of all, a pipeline designed for such high throughput goal would create the infrastructure to enable re-processing of the collection if additional data modalities (e.g. x-ray, infrared, DNA, etc.) become desirable and practical. Furthermore, if the system were sufficiently compact, it could be redeployed for digitization campaigns at a series of institutions.

Using one year as the target timescale, it is possible to estimate parameters that will constrain design choices. For example, if we assume that 50 weeks of 8 hour days is available for digitizing 4.5 million specimens, we must on average process each specimen in about 1.6 seconds on average. This estimate sets the approximate scale of the digitization speed required. Note that allowing five years (7.5 seconds per specimen) for the process doesn't change the overall impression: the pipeline must be able to process specimens very quickly. Existing systems won't capture an object in the time required by our projected needs - no more than a few seconds per object.

An estimate of the amount of image data needed for this digitization process is useful to assess storage and data rate requirements, particularly if either presents a technical challenge [1]. If each of the 4.5 million specimens can be digitized with 12 light field images, each 16 MB in size, then 860 TB will be sufficient for the raw data. This would be enough for the label capture and reconstruction. The data extracted from the labels would require several orders of magnitude fewer bytes to store. Capture of additional modalities, including perhaps very high-resolution reference imagery, will need proportionally more data storage. Existing hardware could handle both the data storage and the data rate needed. This level of data creation, however, is quite significant and it will be advantageous to consider some combination of brute force data compression, selective storage (cropping, grayscale, pruning), or real-time analysis to minimize data storage and transfer hardware costs.

Object size distribution. Knowing the distribution of specimen sizes informs the design of a reliable digitization methodology and the required the capture hardware. With it we can determine the number and kind of digitization stations required in the pipeline. For example, optimal camera placement and sample volume dimensions would depend on this distribution. Specimens large enough to completely obscure labels pinned below them require an entirely different approach to digitization than would very small specimens.

To obtain an initial estimate of this distribution, we pulled 21 drawers at random from the cabinets and took detailed photographic data of the overall composition of each drawer as well as an informal scan of each from several camera angles and positions in order to enable us to assess the general state of the collection from this modest sample. Our analysis is presented in the appendix. The FMNH collection has $93.8 \pm 2.9\%$ specimens smaller than 1 cm wide, $97.6 \pm 2.2\%$ smaller than 2 cm wide, and $99.4 \pm 0.4\%$ smaller than 3 cm wide. The pin and labels, without the insect, fit well in a 2 cm x 2 cm footprint and rise less than 3 cm above the unit box foam base. This dimensional data suggests that most of the collection could be scanned by an instrument with a working volume of about 3 cm cubed. With nominal spacing between specimen and labels, we can expect that as much as 97% and certainly 93% of the collection will present mainly unobstructed views of specimen and labels.

TABLE I
TIME BUDGET (IN SECONDS) PER SPECIMEN IN EACH TIER.

Category	Number of specimens		Time per specimen	
	Minimum	*Maximum*	*Less than*	*At least*
Tier 1 (< 1 cm)	4.1e6	4.3e6	1.8	1.7
Tier 2 (< 2 cm)	90e3	250e3	80	29
Tier 3 (< 3 cm)	4.5e3	160e3	1600	46
Tier 4 (bigger)	9.0e3	45e3	800	160

We use these size categories to create a tentative classification of the specimens into handling categories, or tiers, which can be adjusted as we learn more about the collection. In this scheme, larger specimens are assumed to correlate with more handling or special preprocessing needs, and therefore with longer processing times. The upper and lower bounds estimated for the number of specimens in each category, from 1σ bounds on estimated fraction in each tier, are converted to a

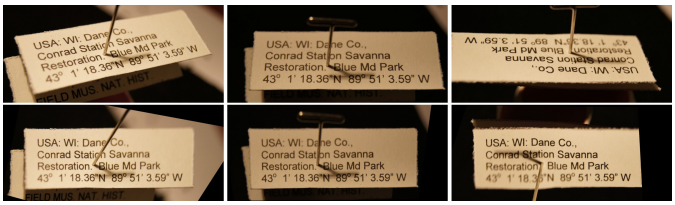

Fig. 1. Preparing multi-view sets of label images for combination to fill in occluded fragments. The bottom image in each column has been reprojected to the image plane to synthesize a rectified view from directly over the label.

time budget for processing each specimen (Table I). We assume that each tier is processed on a dedicated line customized for its special handling needs. The time budget is computed assuming that the line will digitize the entire set of specimens in the tier in one year. The final column reports the time budget assuming the maximum specimen count in that tier. Tegelberg et al. [14] report digitization at 500 pinned insects per day for a workflow that might be appropriate for our Tier 4 handling category. This is comfortably under the 160 seconds per specimen we require for this tier.

Robust label imaging. Because our primary goal is to capture imagery for digitization of the data contained on the associated specimen labels, our hardware and process must ensure that all label surfaces are covered by high quality image data. But the specimens present complex geometry with labels stacked on the pin under the insect, leading to occlusion of label surfaces and no single line of visibility to the label surfaces.

For these reasons we have concentrated on development of multi-view camera heads to capture imagery from many points of view simultaneously. These data can then be analyzed to identify and combine fragments of printed label text. Figure 1 shows an early test using images of a pinned label taken from three widely separated vantage points. In each image the label has been modeled as a simple rectangle in 3-space, which can be reprojected to show the upright label as if viewed from directly above. The pin occludes different parts of the label in each individual image. By combining these three images, as in figure 2, we can create a more

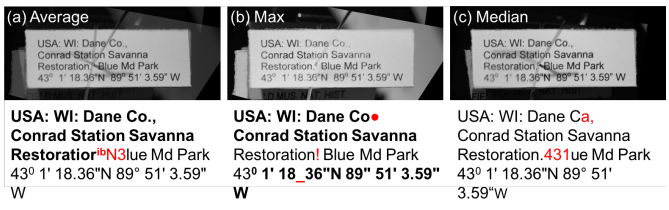

Fig. 2. OCR results for three different combining kernels (a) Average, (b) Max, and (c) Median.

complete label for OCR processing. This simple case illustrates the potential value of our multi-view snapshot approach. In practice, there is often significant occlusion of one label by another or by the specimen, requiring more camera angles and the need to combine more fragmentary glimpses into a coherent whole.

Computational support. Importantly, we distinguish between the computation that must be done in real time and any analyses that may be done offline without compromising the quality of the collected dataset. For this purpose, we note that analysis will be required to aid in provenance tracking, exception detection, and quality assurance. These processes will include automated and continuous tracking of drawers, unit boxes, and specimens.

Much of the analysis can be carried out offline, provided we have imagery of sufficient quality for specimen and label. Algorithms for reducing the raw imagery to fragmentary views of the labels and combining those into virtual replicas of each label will result in a compact database. Analysis of these using OCR, manual transcription, and crowd sourcing methods as necessary will

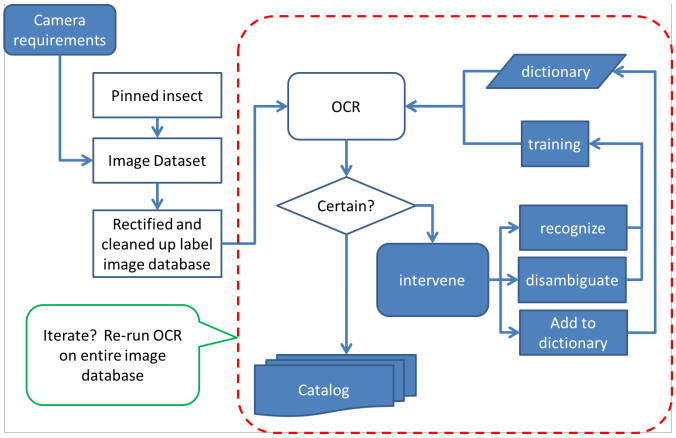

Fig. 3. An adaptive approach to OCR.

create the further compacted data to be included in the final catalog. We remain optimistic that much of this process can be automated and depends only on the quality of the raw data produced during the digitization process. Once captured, the processing required to convert the data into a catalog can be refined and repeated until the results are satisfactory.

The flowchart in figure 3 highlights this important opportunity to improve the accuracy of the automated label reading process. Because the results are not required when the specimen is scanned, the analysis can be repeated while new data is added, enabling evolution of the reader as it is trained on the complex and idiosyncratic properties of the textual data.

Adaptive processing. As the previous discussion shows, many factors will figure into the design and implementation of a successful high-throughput digitization pipeline for pinned insects. In some cases (hopefully most!) imagery can be obtained quickly and without adjustment to the specimen. This would correspond to our Tier 1 handling class in Table I and figure 4. In other cases, label positions and orientations may need to be adjusted to create clear views of all label surfaces (Tiers 2 and 3). In extreme cases, a human may need to remove the labels for imaging (Tier 4). This range of specimen handling needs will require a multi-faceted process that can be adapted to the different cases as they are encountered.

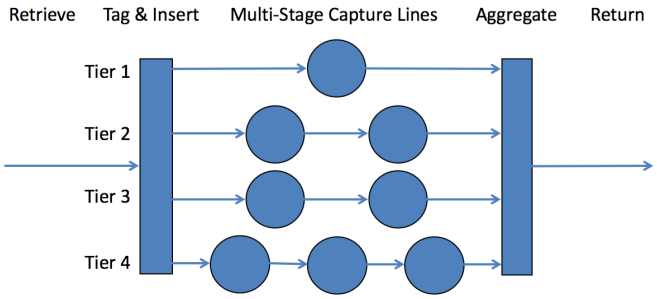

Fig. 4. Schematic of a parallel pipelined processing line for high-throughput digitization.

Figure 4 shows the architecture of a parallel pipelined processing line for this task. Each of the blue circles represents a pre-processing step or a digitization instrument, such as the digitization head we are designing (see figure 6). If the cost per head is low enough, we can increase the throughput by adding lines in parallel. If data must be collected in multiple modalities that cannot be supported by a single head, we can pipeline each line to process incompatible modalities in series without introducing reduction in throughput. The following itemized breakdown describes each phase of the processing pipeline.

- Retrieve. In this step of the process, 60 drawers are staged from the cabinets to the digitization line at any time during the previous day of operation. They will be subject to full drawer imaging, parsing, and registration to initialize the data required for provenance tracking through the system.
- Tag and Insert. Each specimen is given a unique ID, registered with the system, and inserted into the pipeline. The process is computer-guided to ensure insertion into the correct pipeline and to support continued tracking of the objects.
- Multi-Stage Digitization Line. We envision these to be as automatic as is possible. Each parallel line added can increase the overall throughput for a single object class – two lines dedicated to specimens smaller than 1 cm in size would be twice as fast as a single such line. Additionally, lines can be used to target object classes with different characteristics (size, handling requirements) – one line for larger objects, for example.
- Aggregate. At this stage, objects are reconstituted into appropriate unit boxes and then into drawers, guided by a computer process with access to the tracking data.

Missing from this description are the accommodations for exception handling: insect missing from pin, specimen breakage, occlusion identified in real time, specimen too large for available pipeline processing. Exception identification and processing will be integrated with the tracking system.

B. Modular 4π multi-camera 3D capture rig

In developing prototypes of a multi-camera head we aim for single-shot digital capture of casually aligned pinned insects. In order to reliably capture all the information from the labels, the head design must provide for adequate coverage of angle, focus and scale in the face of uncertainty in the pin stack configuration. The resolution required is first set by the demands of OCR.

Using light field cameras in these rigs allows us to sidestep consideration of focal depth, autofocus, and multi-shot focal coverage. The target volume is covered by the collection of rays sampled by the camera which are available for post processing to infinite focus or other optimized applications. For the purposes of prototyping we are using commodity devices, the first generation Lytro light field cameras, because of their size and low cost. Having been discontinued, they are available now for a fraction of their original cost. Suitable high performance compact light field cameras are available from vendors serving the scientific and industrial applications community. The Raytrix R42 is one good example. These will be attractive after prototyping yields a robust configuration for the problem.

Figure 5 shows an example of a multi-camera snapshot produced by a simple 3-camera system (not pictured) with side-by-side cameras arrayed tightly around a central point. This configuration provides a dense array of rays sampling the light field, which for some problems can simplify the 3D reconstruction process.

Our designs also explore ways to position several to many cameras so as to cover the sample volume from many directions. The aim is to provide for views of every portion of the labels despite unknown object orientation

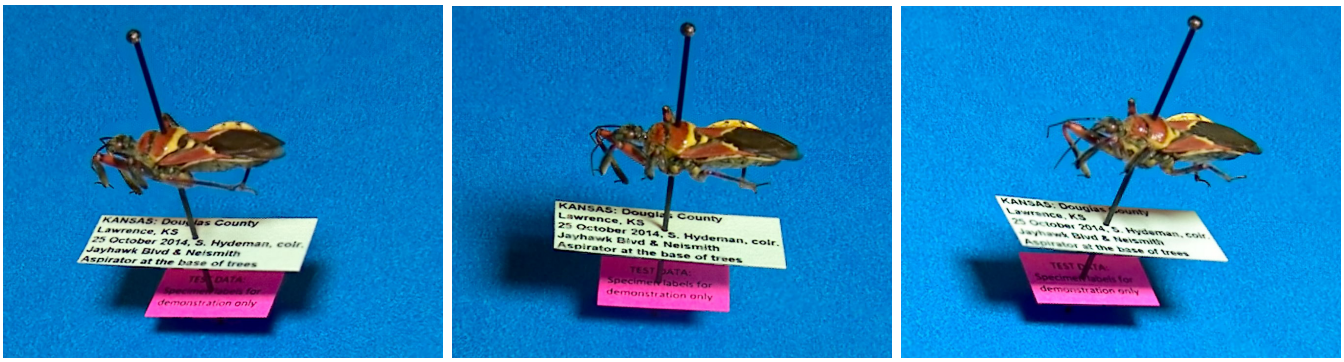

Fig. 5. Example of multi-view capture of a pinned insect.

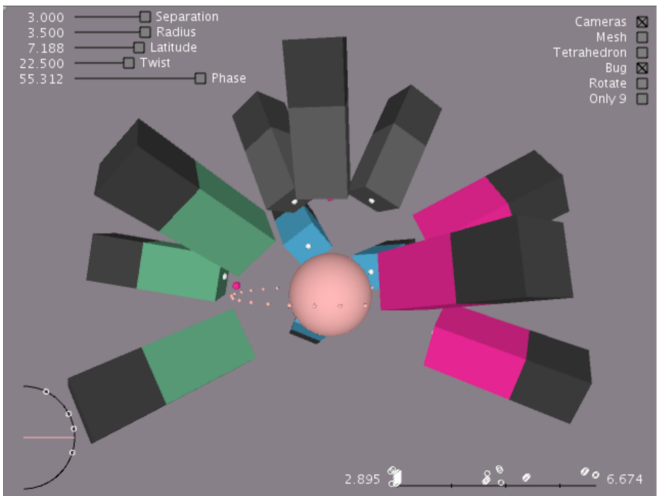

Fig. 6. Design sketch of the modular 4π multi-camera 3D capture rig. Three modules of three cameras, shown in color groups, are arranged around the equator of the sample volume. A single polar module captures the view from the top.

Fig. 7. One of the camera modules surrounded by common objects to convey scale.

and occluded directions. A proof-of-concept camera head is comprised of twelve light field cameras arranged to provide optimal coverage of pinned insects and the data printed on small labels accompanying them on the pin. Figure 6 shows a sketch of this configuration and figure 7 shows one of four modules implemented for this design. With this design we can capture an entire multi-view snapshot in roughly 2 seconds, limited by mechanical actuation of the shutter release button of the commodity light field cameras. Cameras designed for industrial use will capture many images per second.

III. Preliminary Image Analysis

A fully automated capture pipeline requires a number of computational tools as outlined above (section II), including a number of stages that require some form of image analysis. Our preliminary work includes significant progress in capturing text from multiple views of labels, along with supporting methods for capturing high level information about drawer layout. We have begun exploring 3D capture using the light-field camera platform.

Previous work on extracting lines of text from curved documents [6], [8], [9], [13] focuses on fields of text containing many lines and many characters. These techniques were developed for extracting text from camera images of books and generally require a large number of text lines for accurate results. Our labels have only a few lines of text. Additionally, our images are usually not complete labels but have occlusions from other labels and/or the specimen (figure 8).

The approach we are developing identifies text in

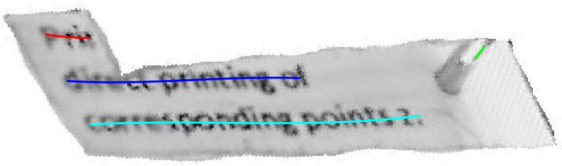

Fig. 8. Our custom label processing must be able to handle label fragments due to occlusion from the specimen or other labels.

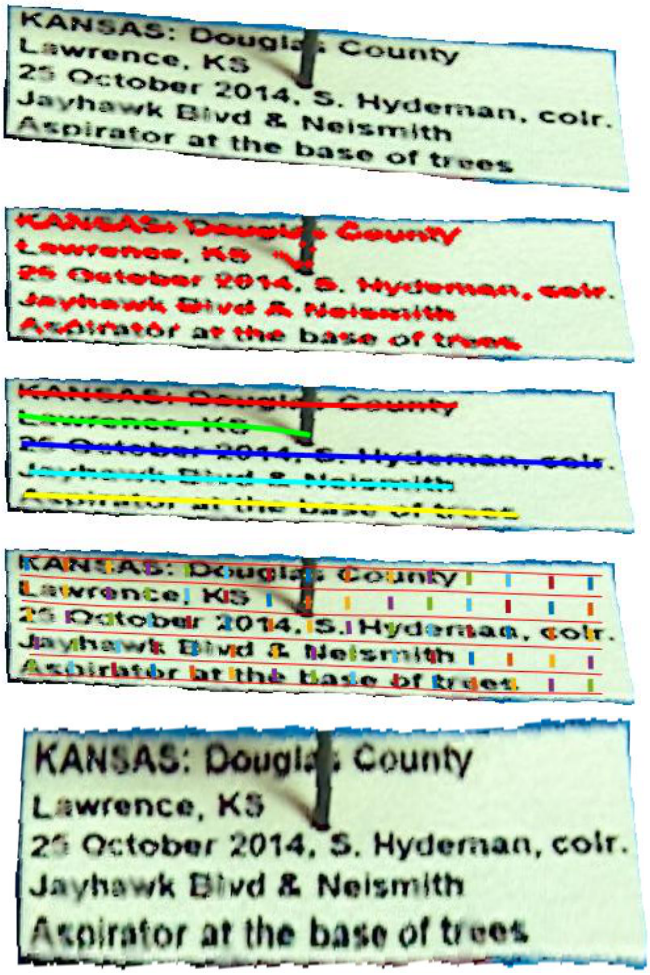

Fig. 9. Processing steps to rectify label text in preparation for OCR.

segmented label fragments and re-projects the fragment. Figure 9 shows key steps in the analysis of each image to extract text from the labels. Character locations are found in each label fragment image and a clustering approach was developed to group these locations into curves, one for each line of text in the label fragment. Using these curves (which are assumed to be projections of lines that are parallel in real space), a grid is positioned on the label image and used to re-project the image into a rectilinear (i.e. fronto-parallel) view. Performing this rectification for all views of each label fragment makes alignment of the fragments simple (e.g. a cross correlation operation). Once rectified and aligned the combined fragments produce sufficient coverage of the label to produce a composite image to be read by OCR (see figure 2).

We also have developed automated methods for drawer level analysis to identify unit trays within the drawer (figure 10). The drawer is found in the image and its corner coordinates are used to create a fronto-planar view. We then use a template-based search to find unit trays of given standard sizes within the drawer. The irregular black border in the figure is an artifact of the projection to rectify the drawer from a slightly off-zenith camera pose. The yellow annotations indicate the identification of several 4x1 trays in the drawer. This information will be used for tracking specimens during drawer processing. Drawer level information (location and size of trays) will be used to guide the processing pipeline and ensure that the drawers are correctly re-populated after specimen digitization.

Our multi-camera system supports 3D capture. There are a number of multi-view 3D reconstruction methods including stereo-based methods provided in open source tools such as openCV [11]. When using light field cameras, each capture provides "multiple" images (one from each lenslet) and 3D information is available for each image. However the depth information from a single light field image provides low resolution 3D information. Using multiple viewpoints allows us to use existing structure from motion techniques. We have experimented with OSM-Bundler [12] and VisualSFM [15] combined with Patch-based/Clustering Views for Multi-view Stereo Software [4], [5] to perform point cloud reconstruction. Initial results are promising. In figure 11 the 3D point cloud has been rendered with each point taking on the color of the ray-stopping pixel, and is set against an artificial blue background in the 3D viewer. Further refinements are needed to fully adapt these methods for light field cameras.

IV. CONCLUSIONS

In this paper we report on design and prototyping of hardware and software to address the problem of digitizing very large collections of pinned insects. We discuss the feasibility of digitizing 4.5 million specimens over the course of one or two years.

Fig. 10. Drawer parsing for specimen tracking and provenance.

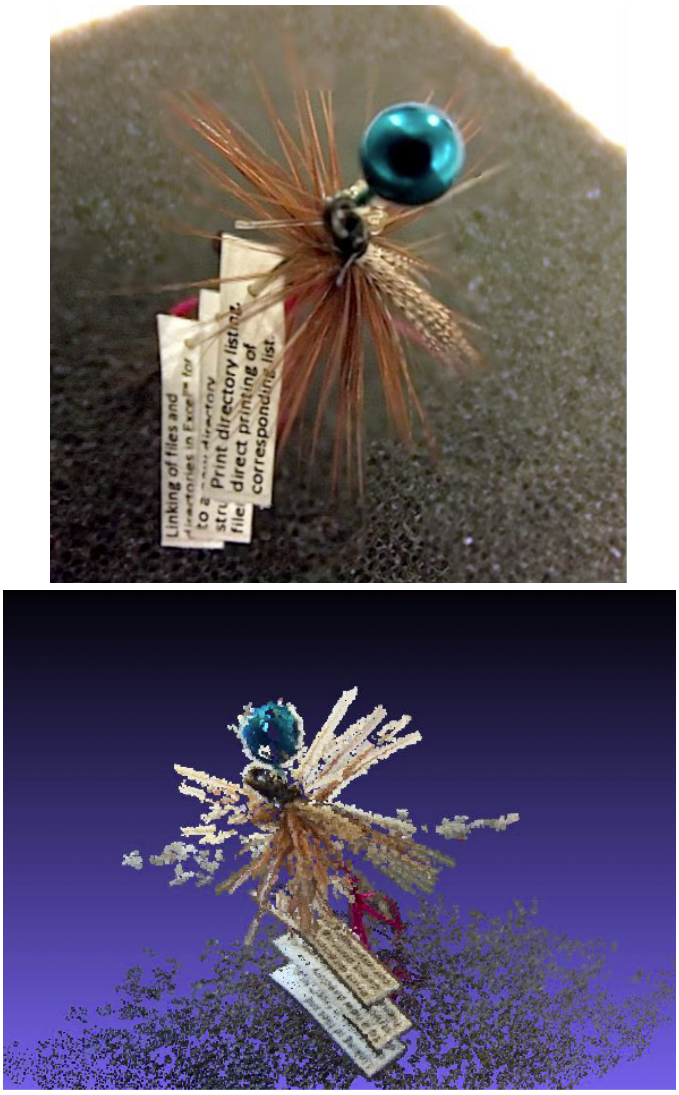

Fig. 11. One view of an experimental pinned specimen (fishing lure) and (bottom) a 3D point cloud reconstruction computed using VisualSFM software.

We present a statistical analysis of the distribution of specimen widths in the collection, which helps to determine how many objects in the collection will require special accommodations. We sketched a multi-stage digitization pipeline and began fleshing out the characteristics of the components by developing designs targeting the bulk of the collection: those objects whose width is small enough to provide line of sight access to the label surfaces.

Our prototype designs provide a platform for testing image analysis algorithms. Snapshot capability enables rapid digital capture in a single exposure. The multi-camera prototype configurations explore pose coverage and resilience against unknown object position and orientation. Light field imaging enables expanded volume coverage without the time required to focus or the need for additional fixed-focus cameras to cover volume.

APPENDIX

Given a process for sampling insects from the larger collection, one can estimate the characteristics of the population. We randomly pulled 21 drawers for inclusion in our sample. We can then calculate the a cumulative distribution of insect size that can be used to inform our system design.

Specimens drawn are unbiased sample

A naive assumption might be to consider the collection of insects in the 21 drawer sample as representative of the 4.5 million specimen population, as if the specimens and not the drawers were drawn randomly from the population. This model ignores biases that may affect the way that specimens are grouped in drawers, instead assuming that each drawer is a collection of randomly selected insects.

Now, because a large specimen will obscure our view of the labels riding below it on the pin, we want to estimate how many specimens have widths less than some limit. The first two columns of Table II give the number of specimens in each of four size bins. The first bin includes insects whose width is no larger than 1 cm; the second bin includes insects whose width is larger than 1 cm but no larger than 2 cm. Assuming, for example, that the probability of a specimen drawn from the population is less than 2 cm wide is p, then the probability that our draw of N from that population includes k specimens less than this 2 cm limit is given by the binomial distribution:

$$P(k,p) = \binom{N}{k} * p^k * (1-p)^{N-k}. \qquad (1)$$

In this example, $N = 5504$ is the number of specimens assumed to be an unbiased draw from the parent population, k is the number of specimens in the draw with a given property (size less than 2 cm, for example), and p is the probability of drawing a specimen from the parent distribution with that property. The distribution of $P(k,p)$ against p provides an estimate for the uncertainty of k as a representation of the expected value $p*N$. The binomial distribution P(k,p) has mean $k = p*N$ and standard deviation $\sigma = sqrt(N*p*(1-p))$. Measured counts and estimates of p and σ are summarized in Table II for this dataset.

Though the drawers are drawn randomly from the entire collection of drawers, some 15,000 in total, there are many ways in which the collection of insects in this subset are correlated. For example, insects of the

TABLE II
AGGREGATE FREQUENCY OF SIZES

Max Size	Count	k	p	σ
1 cm	5178	5178	0.940	0.003
2 cm	203	5381	0.978	0.002
3 cm	94	5475	0.995	0.001
Bigger	29	5504	1.000	0.000

same or related species are typically aggregated in a drawer. Table III shows the breakdown of each of the 21 drawers into size bins. As is evident from the rows of this tabulation, the number of specimens and their distribution differ widely from drawer to drawer.

TABLE III
SIZE SPECTRUM OF EACH DRAWER

Image	1 cm	2 cm	3 cm	Bigger	Total
00952	268	0	0	0	268
00958	0	0	0	11	11
00968	100	0	0	0	100
00990	303	0	0	0	303
01033	170	15	0	0	185
01051	1	46	4	0	51
01070	424	1	0	0	425
01086	311	62	0	0	373
01127	195	0	0	0	195
01149	577	0	0	1	578
01188	317	0	0	0	317
01221	231	0	0	0	231
01264	495	0	0	0	495
01290	249	0	0	0	249
01312	116	39	0	0	155
01333	76	40	0	0	116
01353	196	0	0	0	196
01363	345	0	0	0	345
01375	349	0	0	0	349
01395	455	0	0	0	455
01436	0	0	90	17	107
Totals	5178	203	94	29	5504

Drawers drawn are unbiased sample

The approach we take here is intended to account for these correlated specimen properties within drawers. We consider each drawer to be randomly selected from a population of drawers. For example, we might compute statistics within the drawer: number of specimens, mean size, max size, and mean number of labels per pin. Then each of these might be used in an analysis based on 21 draws from a 15,000 drawer population.

The drawer properties of interest at the moment are derived from the rows in Table III. Normalizing columns 2 ("1 cm") through 4 ("3 cm") by the total in column 6 produces a 3-vector (f_1, f_2, f_3) representing the fraction of the drawer in each of the first three size bins. The sum of the three components of this vector is greater than zero and less than or equal to 1, or $0 <= f_1 + f_2 + f_3 <= 1$.

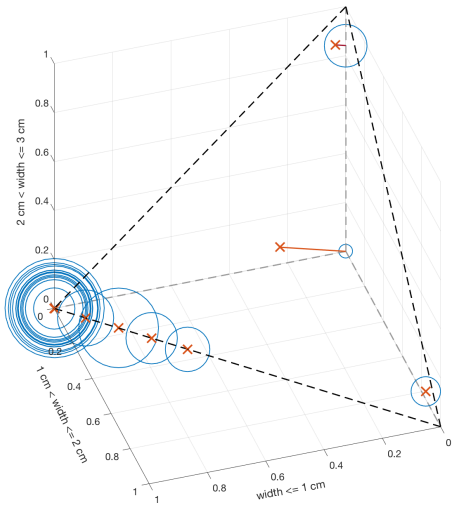

Fig. 12. Distribution of 21 drawers randomly drawn from the population. Each draw is plotted as a circle centered on 3 dimensional coordinates using fraction of specimens in three size bins. The sum of these coordinates is limited to a pyramidal volume between the origin and the triangle $x + y + z = 1$. The size of the circle is proportional to the number of samples in the drawer. A red line segment connects each sample to the perpendicular projection to this triangle as a visual aid.

Figure 12 summarizes this representation of the 21 drawers by plotting each of the size distribution vectors with a circle whose diameter is proportional to the number of specimens in the drawer. The constraint on the sum of the vector components constrain the points to the volume bounded by the dashed lines. The size of the circle is limited (in a complex way) by the requirement that the specimens fit within the area of the drawer.

Even this course sampling of the collection illustrates the correlated size distributions implicit in the organization of the specimens into drawers. Only seven of the 21 drawers include specimens wider than 1 cm. The remaining 14 are centered at (1,0,0). Only one of the drawers (number 00958) is close to the origin, and so is dominated by specimens that are bigger than 3 cm wide. Note also that there are very few specimens in the drawer, because they are relatively large – butterflies, not surprisingly. The rest of the drawers are plotted on or near the unit sum triangle, indicating that nearly all of the specimens in those drawers are less than 3 cm wide.

With many more drawers in our sample, we might construct a smooth representation of their distribution in this space. With the available sample, we can derive an estimate of the size distribution in the parent population as follows. We repeatedly draw a set of M drawers from

the list and compute the cumulative distribution function (CDF) of sizes from each of these experiments. The mean and standard deviation in each bin of the CDF gives us the estimate summarized in figure 13.

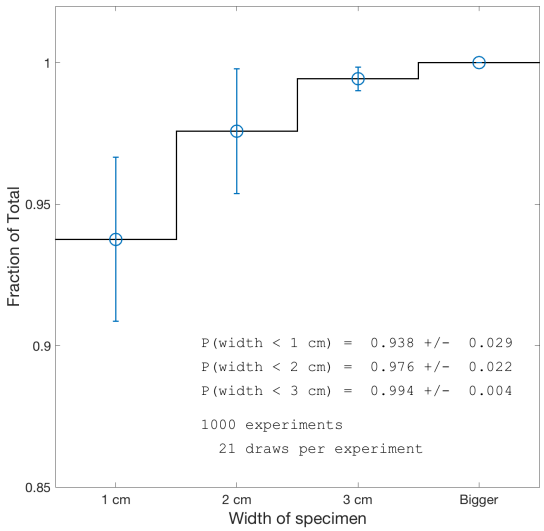

Fig. 13. Cumulative distribution of specimen sizes estimated from a random sample of drawers taken from the FMNH pinned insect collection.

From these analyses we conclude that the FMNH collection has $93.8 \pm 2.9\%$ specimens smaller than 1 cm wide, $97.6 \pm 2.2\%$ smaller than 2 cm wide, and $99.4 \pm 0.4\%$ smaller than 3 cm wide. The means found here are essentially the same as for the initial analysis, as might be expected. The uncertainty is considerably larger, though, as this analysis includes some account of the correlations within drawers.

ACKNOWLEDGMENT

The authors thank Robin De La Pena, Crystal Maeir, and Rudiger Bieler for helpful discussions. Student research interns provided preliminary experimentation during initial project conceptualization (Bryan Dalle-Molle, Julianna Kim, Joshua Koblick, Daniel Shiff, and Andi Zang). MH contributed to experiment design, hardware design, hardware construction, data collection, data analysis, algorithm design, and manuscript writing; NJF contributed to experiment design, hardware design, data collection, data analysis, algorithm design, and manuscript writing; NA contributed to data collection, data analysis, algorithm design, and manuscript writing. PS contributed to experiment design, data collection, and manuscript writing. This material is based upon work supported by the U.S. Department of Energy, Office of Science, under contract number DE-AC02-06CH11357. We acknowledge support for this work from the Field Museum of Natural History.

REFERENCES

[1] Ariño A., Approaches to estimating the universe of natural history collections data, Biodiversity Informatics 7(2): 8192, 2010.

[2] Chapman, A.D., Uses of Primary Species-Occurrence Data, Version 1.0. Report for the Global Biodiversity Information Facility, Copenhagen, 2005. Available at http://www.gbif.org/orc/?doc_id=1300

[3] Duckworth W.D., Genoways H.H., Rose C.L., Preserving natural science collections: chronicle of our environmental heritage, Washington, D.C. iii+140 pp., 1993. Available at www.heritagepreservation.org/catalog/product.asp?intProdID=18

[4] Furukawa, Y., and Ponce, J. (2010, July 13). PMVS2. Retrieved July 12, 2015, from http://www.di.ens.fr/pmvs/

[5] Furukawa, Y. (2010, July 17). CMVS. Retrieved July 12, 2015, from http://www.di.ens.fr/cmvs/

[6] Zuming Huang, Jie Gu, Gaofeng Meng, and Chunhong Pan. Text line extraction of curved document images using hybrid metric. *Proceedings - 3rd IAPR Asian Conference on Pattern Recognition, ACPR 2015*, (95):251–255, 2016.

[7] Hudson LN, Blagoderov V, Heaton A, Holtzhausen P, Livermore L, et al. (2015) Inselect: Automating the Digitization of Natural History Collections. PLOS ONE, 10 (11) : e0143402 - e0143402. doi: 10.1371/journal.pone.0143402

[8] Jian Liang, Daniel DeMenthon, and David Doermann. Flattening curved documents in images. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:338–345, 2005.

[9] Gaofeng Meng, Zuming Huang, Yonghong Song, Shiming Xiang, and Chunhong Pan. Extraction of virtual baselines from distorted document images using curvilinear projection. *Proceedings of the IEEE International Conference on Computer Vision*, 11-18-December-2015:3925–3933, 2016.

[10] Chuong V. Nguyen, David R. Lovell, Matt Adcock, and John La Salle. Capturing natural-colour 3D models of insects for species discovery and diagnostics. *PLoS ONE*, 9(4):1–11, 2014.

[11] http://opencv.org

[12] OSM-Bundler - 3D geometry reconstruction for OpenStreetMap. (n.d.). Retrieved July 11, 2015, from https://code.google.com/p/osm-bundler/

[13] Yuandong Tian and Srinivasa G Narasimhan. Rectification and 3D reconstruction of curved document images. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 377–384, 2011.

[14] Riitta Tegelberg, Tero Mononen, and Hannu Saarenmaa. High-performance digitization of natural history collections: Automated imaging lines for herbarium and insect specimens. *Taxon* 63.6 (2014): 1307-1313.

[15] VisualSFM: A Visual Structure from Motion System. Retrieved July 11, 2015, from http://ccwu.me/vsfm/