3D surface reconstruction of entomological specimens from uniform Multi-view image datasets

Germán D. Sosa*[1], Sebastián Rodríguez[1], Javier Guaje[2], Jorge Victorino[1], Manuel Mejía[1], Luz Stella Fuentes[3], Angélica Ramírez[1], and Hugo Franco*[1]

[1]Universidad Central, [2]Universidad Nacional de Colombia, [3]Universidad Jorge Tadeo Lozano
[1,2,3]Bogotá, Colombia
*e-mail: {gsosar,hfrancot}@ucentral.edu.co

Abstract

Modeling of 3D objects and scenes have become a common tool in different applied fields from simulation–based design in high–end engineering applications (aviation, civil structures, engine components, etc.) to entertainment (computer-based animation, video-game development, etc.). In Biology and related fields, 3D object modeling and reconstruction provide valuable tools to support the visualization, comparison and even morphometric analysis in both academical and applied tasks.

Such computational tools, usually implemented as web-based virtual reality applications, significantly reduce the manipulation of fragile samples, preventing their damage and, even, their complete loss. On the other hand, they allow to take the morphological properties of physical specimens to the digital domain, giving support to common entomology tasks such as characterization, morphological taxonomy and teaching. This paper addresses the problem of producing reliable 3D point clouds from the surface of entomological specimens, based on a proved approach for multiview 3D reconstruction from high resolution pictures. Given the traditional issues of macro-photography for small sized objects (i.e. short depth of field, presence of subtle and complex structures, etc.), a pre-processing protocol, based on focus stacking, supported the generation of enhanced views obtained by an acquisition device specifically designed for this work.

The proposed approach has been tested on a sample of six representative subjects from the Entomological Collection of the Centro de Biosistemas, Universidad Jorge Tadeo Lozano (Colombia). The resulting point clouds exhibit an overall good visual quality for the body structure the selected specimens, while file sizes are portable enough to support web based visualization.

1. Introduction

Over the last few decades, computer generated 3D models have supported variety of fields such as medicine, engineering, science, etc. In particular, biological applications have greatly benefited from such advanced visualization systems, which implement valuable tools to improve and ease human-dependent tasks s.a. annotation, characterization and modeling of biological specimens. Given the fast evolution of the available computational power and the formulation of robust yet efficient computer vision algorithms, 3D modeling from images shows itself as a cost-effective alternative to sophisticated systems like laser-based 3D scanning [16], Micro-Computer tomography (μ-CT)[8, 9], microscope scanning [12, 15] or ultra-microscopy [5]. Indeed, 3D-reconstruction methods based on image sequences have proved to be an affordable option with remarkable results for textured and high quality 3D models.

Most of the state-of-art related approaches have grouped around two methodologies: Visual Hull (VH) [18, 19] a.k.a Shape From Silhouette (SPS) algorithm, and Multiview Stereo (MVS) [23]. While VH is a common option to address 3D reconstruction due to its easy implementation [3], it does not take into account color nor texture content of the object, rather than its silhouette, making it prone to errors when occlusions or concave regions are present [14]. On the other hand, multiview stereo uses a measure of photo-consistency, based on the image texture or color content, exhibiting pretty good results even using fewer views of the object than standard VH approaches [14]. Nevertheless, VH methods are still frequently used to initialize MVS implementations, since they are usually robust to outliers [6, 10].

Recent works in this field have applied several different multiview approaches to perform 3D reconstruction on moderate–sized images (1–4 megapixels) [10, 14, 16, 17], since they apply highly demanding techniques in terms of computational cost and memory requirements, s.a. surface or volumetric optimization or VH constraints, not suitable

978-1-5090-3797-1/16/$31.00

for larger images. Given that, the main goal of this work is to propose a systematic method to generate reliable and dense point clouds representing the surface of preserved entomological samples based on multiple high-resolution images. Such virtual representations of physical specimens could support time–consuming human–driven tasks like characterization, morphological exploration, morphometry and certain aspects of morphological taxonomy in a non-destructive approach. This scenario implies the implementation of a MVS 3D surface reconstruction method for small sized objects, addressing the presence of insect complex and subtle structures –antennas, wings, legs, etc.–, concave surfaces and reflective exoskeleton regions [21]. However, macro-photography involves issues like short depth of field, which demands post-acquisition techniques like Focus Stacking [4], in order to produce high quality textured images able to provide good enough results via MVS methods. To address this requirement, an automated tool and a image acquisition protocol similar to those described by Nguyen et.al. [21] and Gallo et.al.[11]. Once full image sequence is obtained, a computational algorithm based on Tola et.al.[25] approach was implemented to get dense pointclouds ready to be rendered into a realistic 3D model of a entomological specimen.

2. Materials and Methods

The method proposed in this work comprises three main stages: i) an automated image acquisition protocol to capture high resolution images of a subject from different points of view; ii) a post-processing method involving Focus Stacking and Background Suppression to obtain full calibrated image sequences; and iii) a computational method based on a MVS approach suited for high resolution images is implemented to get realistic point clouds.

2.1. Multiview image acquisition

For image acquisition and 3D reconstruction, a representative subset of subjects from the Entomological Collection of the *Centro de Biosistemas*[1] were selected according to these criteria: i) Well preserved specimens of different orders with no damage on its external structure and able to hold its original shape during manipulation in the image acquisition process. ii) Every sample must to fit into $1.5 \times 2.5 \times 2.5$ cm box to avoid out-of-view issues when taking photographs. iii) Every sample must be attached to a metallic pin to hold straight alignment while were gripped by the acquisition device. Using these criteria, a group of 6 different specimens were selected.

[1]Part of the Universidad Jorge Tadeo Lozano (Colombia), located in the town of Chia (26 km to the north of Bogotá). The complete collection comprises more than 10,000 subjects, including samples of different orders (e.g. Coleoptera, Hymenoptera, Diptera, Lepidotera, etc.)

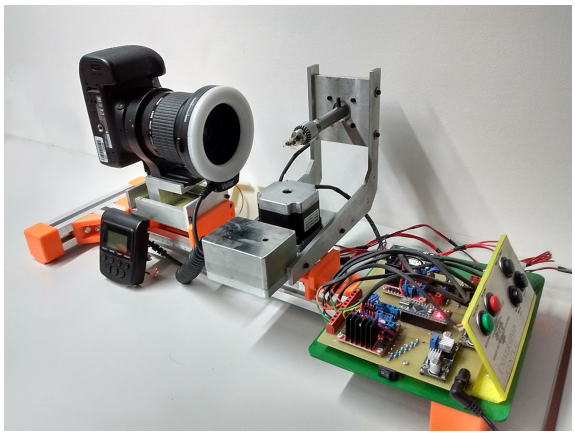

Figure 1. Image acquisition device.

2.1.1 Acquisition device and protocol

Based on the automated acquisition protocol described by Nguyen et.al. [21][20] we performed a similar approach for image acquisition from *uniformly* spaced points of view (POV) (all pictures were taken at the same distance from subject and equally spaced from each other). To acquire such uniform image sequence, a specific device was designed in the *Universidad Central (Colombia)* to support this application. This device consists of two stepper motors, a Canon EOS 70D digital camera with a Canon EF-65 mm macro lens and a manually adjustable rail with a micrometric screw (Figure 1).

The stepper motors provide two angles of freedom (tilt and pan) and allow to rotate the subject in a wide range of angular positions. A full sequence of images is obtained as follows: The first stepper motor performs tilt movement and turns the subject in three specific elevation angles : $-45°$, $0°$ and $45°$ respectively as shown in Figure 3 (a); the points of view induced for such movement correspond to those shown in (b). For each elevation angle, the second motor performs a pan movement and rotates the subject $360°$automatically shooting the camera every $18°$as shown in (c); the points of view induced here correspond to those shown in (d). Combining both movements, this device is able to take $20 \times 3 = 60$ photographs producing an uniform point of view pattern (Fig. 4).

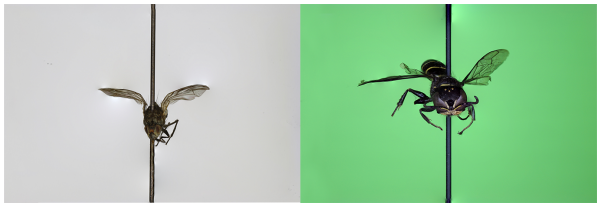

Figure 2. Two entomological samples used to evaluate the proposed method: Diptera (fly, left) and Hymenoptera (Wasp, right).

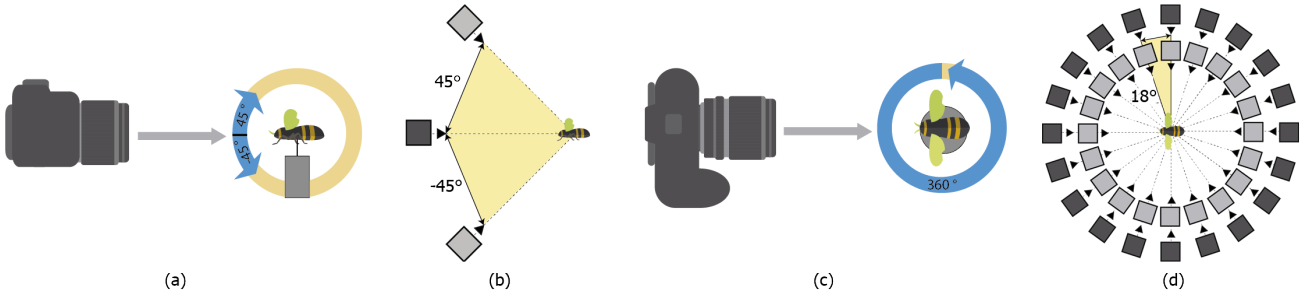

Figure 3. Illustration of image acquisition protocol. Tilt movement makes a rotation of the specimen within in the median plane for three different angle positions: 45°, 0°and -45°(a) generating different visual orientations for picture acquisition: above, inline and below the subject respectively (b). The pan movement turns the specimen 360°in the frontal plane (c) allowing to take pictures of subject from variety of angles around it(d).

Despite the multiple POV achieved by the device can cover almost every detail from the specimen, the short depth of field of the Canon macro lens (as well as other macro lenses) focus only a thin area of the subject at each POV. Since EF-65 lens has a fixed focal length, the camera must be displaced along the direction of specimen to focus different sections. The rail then provides a third degree of freedom by moving the camera along in the axial direction at micrometric steps. Then, a complete sequence of 60 photos is carried out by the two stepper motors and automatic remote camera shooter; when whole sequence is completed, camera is shifted 1 mm in the direction of subject and a new sequence of 60 photos is acquired. This process is repeated until camera focus reach the most rear section of subject, so for a insect of 15 mm long, device has to take $15 \times 3 \times 20 = 900$ photos to retrieve all texture and color information of specimen

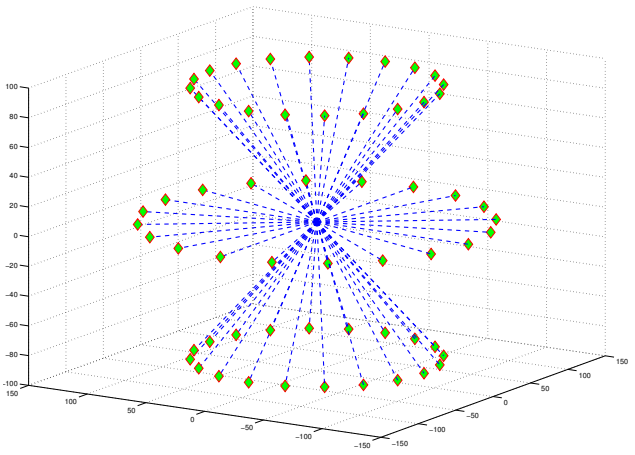

Figure 4. Point of view pattern used by the image acquisition protocol (Figure 3). Diamonds show each camera position while blue dotted lines represent the camera orientation as it faces the rotation center of the acquisition device (corresponding to the subject centroid).

2.1.2 Camera calibration

To approximate the 3D shape of an object given a set of 2D pictures, it is necessary to know how the 3D object is projected into a 2D space for each one of those images using *central projection* [13], which is a linear transformation in homogeneous coordinates between a 3D point in the real world, X, and a 2D point projected on each picture, x – Eq. (1). The 3 coordinates of X are then projected into the two components of x through a matrix projection, P. This projection matrix contains information about the geometric model of the camera and how it is oriented with respect to the real world. The z and Z are the additional components in the homogeneous coordinate representation of X, x points, and allow to find their original Cartesian representation.

$$[x_1, x_2, z]^T = P_{3\times4}[X_1, X_2, X_3, Z]^T \tag{1}$$

Equation (2) shows how projection matrix (a.k.a fundamental matrix) can be decomposed into 3 separated components: a set of intrinsic parameters K related to the geometric pinhole camera model representation, a rotation matrix R that defines the change from real–world coordinates to camera coordinates along with the orientation of camera with respect to scene, and a translation vector $t = -RC$ describing the translation of camera center around the subject given the camera center coordinates C.

$$P_{3\times4} = K_{3\times3}[R_{3\times3}|t_{3\times1}] \tag{2}$$

Knowing the projection matrix P is crucial to infer 3D points locations from different sets of two-dimensional projected points, however, the estimation of this matrix is not trivial and is known as the *pose estimation* problem that normally requires third-party optimization algorithms as well as a pattern calibration image [24]. Nonetheless, given the accurate movement of stepper motors in conjunction with the Canon camera and lens specifications we are able to provide a good *a priori* estimation of such parameters based on

the formulation for each component of P given by the following set of formulas in Equation (3).

$$R = R_{wc}\begin{bmatrix}1 & 0 & 0\\ 0 & cos(\phi) & -sin(\phi)\\ 0 & sin(\phi) & cos\phi\end{bmatrix}\begin{bmatrix}cos(\theta) & sin(\theta) & 0\\ -sin(\theta) & cos(\theta) & 0\\ 0 & 0 & 1\end{bmatrix} \quad (3)$$

$$K = \begin{bmatrix}f/\mu & 0 & p_x/2\\ 0 & f/\mu & p_y/2\\ 0 & 0 & 1\end{bmatrix} \quad t = -R\left(d\begin{bmatrix}cos(\phi)cos(\theta)\\ cos(\phi)sin(\theta)\\ sin(\phi)\end{bmatrix}\right)$$

R rotation matrix is considered as an extrinsic parameter that takes the tilt and pan angles, ϕ and θ respectively, to calculate the full orientation of camera in camera coordinates, hence a 3×3 transformation matrix R_{wc} is required to convert from world to camera coordinate system. Another extrinsic parameter is the translation vector t that is calculated based on the value of same tilt and pan angles, ϕ and θ, and the focus distance from subject d. Finally, the intrinsic parameters about the camera–lens geometric model are included in K based on values of focal length f, pixel size μ and resolution in pixels (p_x, py), (Skew and distortion factors were considered negligible for the camera and lens selected). For all points of view in Figure 4, the same configuration of K is used, in contrast of the value of t and R that vary depending on the actual values of ϕ and θ.

2.2. Image pre–processing

Section 2.1. describes the image acquisition process to capture multiview texture and color data from uniform picture sequences of an entomological specimen. Due to the tight depth-of-field associated to macro lenses, the acquisition device should provide a mechanism to obtain several partially–focused pictures for each POV, in order to implement a Focus Stacking algorithm yielding a fully–focused image. Additionally a background suppression process is applied to remove irrelevant points for 3D reconstruction, reducing the computation time.

2.2.1 Focus Stacking

Focus stacking [4] is an image processing technique combining multiple pictures taken at different focus distances to obtain a synthetic image with greater depth-of-field than any of the individual source pictures [22], commonly find in macro photography and optical microscopy applications.

For this particular application, all images for the same point of view are combined using a focus stacking tool included by *Adobe Photoshop* software. Number of photos to combine may vary depending on the size of specimen, for example, a 15mm insect requires 15 photos every millimeter to cover all individual sections. However, regardless of the size of insect, the FS process yields an unique image per point of view; so for a whole full image acquisition sequence FS returns a reduced image collection of $3\times20 = 60$ pictures. Figure 2 shows the results of FS for the same point of view in two different samples.

2.2.2 Background suppression

As explained in section 2.3, MVS algorithms perform a per pixel measure of photoconsistency on different pairs of images to estimate the location of a real 3D point through stereoscopy. Even though, it is not necessary to perform such measure on pixels that do not correspond with the subject of interest (i.e. entomological specimen) but belong to background region.

To reduce the computational cost of MVS implementation, a binary mask was created per image selecting the Region of Interest (ROI) corresponding to insect body where the true logical values correspond to pixels being measured for photoconsistency while negative values belongs to background and are not considered into MVS algorithm.

The actual impact of background suppression process becomes noticeable when dealing with small subjects since most of image turns into background reducing the time of computation in a factor of 2 to 8 depending of the size of the specimen. Additionally, binary masks are later used to make a cleaning operation of outliers in pointclouds since they resemble the silhouettes of subject and can be used to run a VH fashion post-processing stage in a similar to previous approaches for 3D reconstruction from images[3,6,10].

2.3. Multiview–based 3D reconstruction

A novel proposal carried out by Tola et.al [25] describes an MVS approach designed to work with ultra-high resolution images with an efficient cost of memory and computational time. The core of this approach is a robust photoconsistency measure between images based on DAISY descriptors[7]. DAISY are able to capture high texture content of large resolution images with relative low memory and computational cost requirements bringing reliable matches for dense 3D pointclouds generation avoiding the use of posterior optimization constraints.

Tola et.al. reported results motivated the implementation of a similar approximation for this work about 3D reconstruction of entomological specimens based on high resolution and textured images provided by our acquisition tool. Our proposed method can be divided in the following sections: Photoconsistency measures based on DAISY descriptors, photoconsistency enforcement via multiple views, and use of binary masks to perform VH fashion post-processing for outlier suppression.

2.3.1 Photoconsistency measure

The core of every MVS implementation is the photoconsistency measure because allow depth estimation of single

Algorithm 1 3D point estimation algorithm

Input: Image pair (I_i, I_j)
Input: I_i 2D point at m, n coordinates $(x^i_{m,n})$
Input: POV fundamental matrices P_i, P_j
Input: Depth along epipolar line (λ)
Input: Depth step along epipolar line $(d\lambda)$
Input: I_j image resolution (u_{max}, v_{max})
Output: 3D point location for $x^i_{m,n}$ $(X^{i,j}_{m,n})$

procedure COMPUTE3D(I_i,I_j,$x^i_{m,n}$,P_i,P_j)
  d_i = computeDaisy(I_i,$x^i_{m,n}$)
  $x^j_{u,v}$ = epipolarProjection(I_j,$x^i_{m,n}$,P_i,P_j,λ)
  while $u < u_{max}$ and $v < v_{max}$ **do**
    d_j = computeDaisy(I_j,$x^j_{u,v}$)
    D = append(D,d_j,$x^j_{u,v}$)
    $\lambda = \lambda + d\lambda$
    $x^j_{u,v}$ = epipolarProjection(I_j,$x^i_{m,n}$,P_i,P_j,λ)
  end while
  r = bestScoreRatio(d_i,D)
  if $r > 1.25$ **then**
    $x^j_{u,v}$ = bestScorePoint(d,D)
  end if
  $X^{i,i-1}_{m,n}$ = project3D($x^i_{m,n}$,$x^j_{u,v}$,P_i,P_j)
  return $X^{i,j}_{m,n}$
end procedure

points x given a pair of images through stereoscopy. A robust and computationally efficient photoconsistency estimation facilitates the implementation of an efficient 3D reconstruction algorithm in a moderate time for large sequences of images.

Current approach implements photoconsistency based on DAISY descriptors and is presented in Algorithm 1. Given a pair of images I_i, I_j and one (m, n) 2D point in image I_i i.e. $x^i_{m,n}$, a DAISY descriptor d_i is calculated for that point (*computeDaisy*). Then the most photoconsistent point in image I_j must be found to give actual 3D position through stereoscopy. Usually, finding that point requires $\mathcal{O}(n^2)$ computation time since images are two-dimensional, however epipolar geometry reduces this problem to a linear search given the calibration matrices P_i, P_j. The calculation of the epipolar line (*epipolarProjection*) can be consulted in the Tola et.al work [25] in the Appendix.

For every point $x^i_{m,n}$ a number of samples are projected onto the epipolar line by changing the value of depth λ along the line in small steps of $d\lambda$. Every sample generates a 2D point positioned at u, v in image I_j called $x^j_{u,v}$, for each sample algorithm calculates its DAISY descriptor d_j and is appended into an array D along with the position of $x^i_{u,v}$ (*append*). Then a score function –equation 4– is used between d_i and each one of samples on D to determine the most photoconsistent point.

$$S(d) = e^{-\frac{||d_i - d_j||^2}{\sigma}} \tag{4}$$

Where d_i and d_j are the DAISY descriptors of point $x^i_{m,n}$ and sampled points $x^j_{u,v}$ onto the epipolar line, the σ coefficient controls the sharpness of the distribution penalizing or not high differences between descriptors. Score function is used to estimate the most photoconsistent pair of points, however an additional constraint is taking into account to assure the validity of selected pair. The constraint is that if the ratio between first and second best score is greater that 1.25 then it can be trusted that selected pair is valid (*bestScoreRatio*), if not is deprecated.

Finally, if the score function ratio finds a valid pair of points, $x^j_{u,v}$ is assigned with the most photoconsistent sample of D (*bestScoreRatioIndex*). Then both points $x^i_{m,n}, x^j_{u,v}$ in conjunction with calibration matrices P_i, P_j are used to estimate actual 3D location $X^{i,j}_{m,n}$ from images I_i, I_j(*project3D*), as it is shown in [25].

2.3.2 Enforcing consistency

An important observation about photoconsistency is that is performed using pairs of images, however it is possible to exploit information of additional views to refine obtained 3D points for each single image I_i e.g. the method described in Algorithm 2. This refining process takes the obtained 3D point through two subsequent images $X^{i,i-1}_{m,n}$ and compares its relative error (*consistencyError*) against the same calculation of 3D point but using different pairs of images $X^{i,j}_{m,n}$ for $j = \{i-2, i+1, i+2\}$ assuming that selected 2D point $x^i_{m,n}$ can be viewed for all selected pairs.

Relative error e is calculated using the depth along the epipolar line λ for different combinations of 3D points $\{(X^{i,i-1}_{m,n}, X^{i,i-2}_{m,n}), (X^{i,i-1}_{m,n}, X^{i,i+1}_{m,n}), (X^{i,i-1}_{m,n}), X^{i,i+2}_{m,n})\}$. If this error is lower than 5%, then such pair is considered consistent and C is incremented. Finally, only the points X^i are preserved, such that all pair combinations are consistent i.e. $C = 3$

2.3.3 Outlier supression through binary masks

Both previous processes allow the calculation and enforcement of 3D points arrays X^i for an unique image I_i. The complete point cloud is generated by appending the entire group of 60 arrays into a large array of points called X. However, final pointcloud may present some outliers. Thus, we proposed the use of binary masks to perform a refining process based on silhouettes that performs a similar work to visual hull method but are much more cost-efficient in computational terms.

A final cleanup of the point cloud X projects the whole array, using Equation 1, on every binary mask B_i, remind that binary masks have true values for those point in ROI

Algorithm 2 Consistency enforcement algorithm

Input: Image (I_i)
Input: I_i 2D point at m, n coordinates ($x^i_{m,n}$)
Input: POV fundamental matrix P_i
Input: I_i image resolution (m_{max}, n_{max})
Output: 3D estimation of all ROI points in image I_i (X^i)

procedure ENFORCECONSISTENCY(I_i,P_i)
  $C = 0$ ▷ consistent view counter
  $e = 0$ ▷ consistency error
  for m=0 **to** m_{max} **do**
    for n=0 **to** n_{max} **do**
      $X^{i,i-1}_{m,n}$ = compute3D(I_i,I_{i-1},$x^i_{m,n}$,P_i,P_{i-1})
      for j **in** $i-2, i+1, i+2$ **do**
        $X^{i,j}_{m,n}$ = compute3D(I_i,I_j,$x^i_{m,n}$,P_i,P_j)
        e = consistencyError($X^{i,i-1}_{m,n}$,$X^{i,j}_{m,n}$)
        if $e < 0.05$ **then**
          $C = C + 1$
        end if
      end for
      if $C = 3$ **then**
        $X^i_{m,n}$ = $X^{i,i-1}_{m,n}$
        $X^i = append(X^i, X^i_{m,n})$
      end if
    end for
  end for
  return X^i
end procedure

and false for those in the background. If reprojected points fall into ROI for all masks is preserved into the final point cloud as described in Equation 5, otherwise are eliminated.

$$X = \left\{X^i : \forall x^i = P_i X^i, x^i \in ROI(B_i)\right\} \quad (5)$$

3. Results

To evaluate the quality of 3D models obtained by the proposed approach in a real scenario, two representative physical samples were selected, corresponding to Diptera and Hymenoptera (Figure 2). For each specimen, an image sequence of 60 photos from different angles was acquired, according to the acquisition protocol. However, the number of FS pictures per POV varied from 10 for the fly ($\sim$ 1.0 cm long) to 18 for the wasp ($\sim$ 1.8 cm long), yielding a total of $10 \times 3 \times 20 = 600$ and $18 \times 3 \times 20 = 1080$ photos, respectively. All pictures were shot at a resolution of 5MP (2736×1824 pixels) with an average file size of 1.2 MB. The entire collection size is approximately 72 MB after the Focus Stacking process. Resulting point clouds can be observed in Figure 5 and correspond to dense sets of around 1 million points for a fly and 2.4 millions for wasp. The resulting point clouds are available via Web in [1,2], showing that realistic 3D reconstruction for such complex and small sized objects is possible using low–cost acquisition tools (based on digital cameras and macro lenses) in comparison to laser scanners and optical microscopes.

The point clouds obtained by applying the proposed method to the selected samples are shown in Figure 5). Even they evidence an acceptable and promising performance, there are several drawbacks like incomplete regions in wings and legs. We attribute those problems to the automated background suppression process since wings are translucent and take a similar color to background panel , particularly on the green wings of wasp. On the other hand, incomplete legs may be caused due to slight inconsistencies between *a priori* calibration parameters given by 3 and actual positions of the acquisition tool.

Respect to the computational cost, a sequence of 60 photographs was used for each sample to perform 3D reconstruction of specimens, all photographs were taken with resolution of 5MP (2736×1824 pixels). 3D reconstruction algorithm was implemented and executed in MATLAB R2012a on a 2.4 GHz 16-cores Intel Xeon CPU and 16 GB RAM server running Linux Ubuntu 12.04 LTS. Despite the large computational capacity provided by the server, the algorithm only used 1 core for each POV combination (4 in total). Additionally, MATLAB is an interpreted language which performs 10 to 100 times slower (in particular for nested loops) in comparison with compiled programming languages e.g. C++. Given that, the 3D reconstruction process takes $\sim$ 4 hours for the Diptera specimen and $\sim$ 10 hours for the Hymenoptera specimen.

4. Conclusions and future work

This work proposes a complete implementation of a framework intended to support 3D model reconstruction from high-resolution image sequences applied to the digitization of entomological specimens consisting in two main stages: an image acquisition protocol (that integrates macro-photography and Focus Stacking to produce high resolution textured images, based on the work of Nguyen et.al.[21]), and a multiview 3D reconstruction algorithm (based on the work of Tola et.al.) [25], providing 3D dense and realistic point clouds from multiview sequences of high resolution images.

After visual inspection, the results obtained by this method are good enough for a preliminary version. Incoming stages of the development include corrections for background suppression process and camera parameter calibration to avoid incomplete regions in final point clouds. Moving the code to compiled programming languages and improving the algorithm so it can be optimized and parallelized are also under consideration, in order to use stricter parameters in much shorter times.

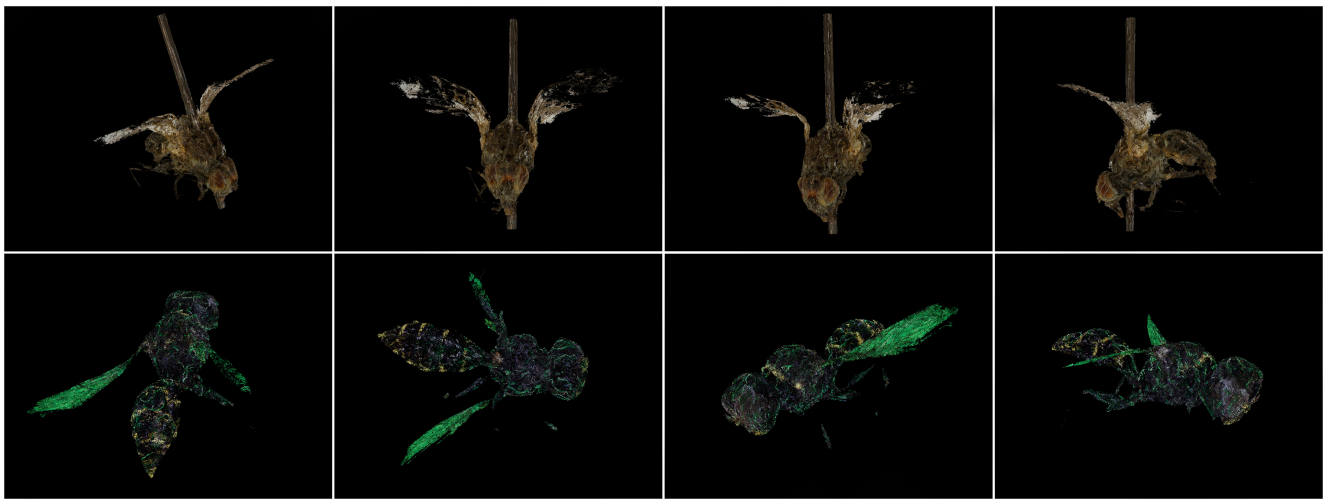

Figure 5. Obtained point clouds for two specimens of entomological collection. A fly (top) and a wasp (bottom)

References

[1] 3D Reconstructed fly, pointcloud visualization. http://www.pointclouds.org/assets/viewer/pcl_viewer.html?load=http://hpclab.ucentral.edu.co/~gsosar/bugFly_full.pcd. Accessed: 2016-06-19.

[2] 3D Reconstructed wasp, pointcloud visualization. http://www.pointclouds.org/assets/viewer/pcl_viewer.html?load=http://hpclab.ucentral.edu.co/~gsosar/bugWasp_full.pcd. Accessed: 2016-06-19.

[3] K. Atsushi, H. Sueyasu, Y. Funayama, and T. Maekawa. System for reconstruction of three-dimensional micro objects from multiple photographic images. *Computer-Aided Design*, 43(8):1045–1055, 2011.

[4] P. Bargh. Focus stacking—a macro photography technique. *You Tube video, retrieved on Dec*, 10, 2013.

[5] K. Becker, N. Jährling, E. Kramer, F. Schnorrer, and H.-U. Dodt. Ultramicroscopy: 3d reconstruction of large microscopical specimens. *Journal of biophotonics*, 1(1):36–42, 2008.

[6] D. Cremers and K. Kolev. Multiview stereo and silhouette consistency via convex functionals over convex domains. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(6):1161–1174, 2011.

[7] V. L. E. Tola and P. Fua. DAISY: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5):815–830, 2010.

[8] S. Faulwetter, A. Vasileiadou, M. Kouratoras, T. Dailianis, and C. Arvanitidis. Micro-computed tomography: Introducing new dimensions to taxonomy. *ZooKeys*, (263):1, 2013.

[9] F. Friedrich and R. G. Beutel. Micro-computer tomography and a renaissance of insect morphology. In *Optical Engineering+ Applications*, pages 70781U–70781U. International Society for Optics and Photonics, 2008.

[10] Y. Furukawa and J. Ponce. Accurate, dense, and robust multiview stereopsis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(8):1362–1376, 2010.

[11] A. Gallo, M. Muzzupappa, and F. Bruno. 3d reconstruction of small sized objects from a sequence of multi-focused images. *Journal of Cultural Heritage*, 15(2):173–182, 2014.

[12] S.-Q. Ge, B. Wipfler, H. Pohl, Y. Hua, A. Ślipiński, X.-K. Yang, and R. G. Beutel. The first complete 3d reconstruction of a spanish fly primary larva (lytta vesicatoria, meloidae, coleoptera). *PloS one*, 7(12):e52511, 2012.

[13] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[14] C. Hernández and G. Vogiatzis. Shape from photographs: a multi-view stereo pipeline. In *Computer Vision*, pages 281–311. Springer, 2010.

[15] A. Klaus, V. Kulasekera, and V. Schawaroch. Three-dimensional visualization of insect morphology using confocal laser scanning microscopy. *Journal of Microscopy*, 212(2):107–121, 2003.

[16] K. Kolev, M. Klodt, T. Brox, and D. Cremers. Continuous global optimization in multiview 3d reconstruction. *International Journal of Computer Vision*, 84(1):80–96, 2009.

[17] K. Kolev, T. Pock, and D. Cremers. Anisotropic minimal surfaces integrating photoconsistency and normal information for multiview stereo. In *Computer Vision–ECCV 2010*, pages 538–551. Springer, 2010.

[18] A. Laurentini. The visual hull concept for silhouette-based image understanding. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(2):150–162, 1994.

[19] W. Matusik, C. Buehler, R. Raskar, S. J. Gortler, and L. McMillan. Image-based visual hulls. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 369–374. ACM Press/Addison-Wesley Publishing Co., 2000.

[20] C. Nguyen, D. Lovell, R. Oberprieler, D. Jennings, M. Adcock, E. Gates-Stuart, and J. La Salle. Natural-color 3d insect models for education, entertainment, biosecurity and

science. In *ACM SIGGRAPH 2014 Posters*, page 107. ACM, 2014.

[21] C. V. Nguyen, D. R. Lovell, M. Adcock, and J. La Salle. Capturing natural-colour 3d models of insects for species discovery and diagnostics. *PloS one*, 9(4):e94346, 2014.

[22] S. F. Ray. *Applied photographic optics: Lenses and optical systems for photography, film, video, electronic and digital imaging*. Focal Press, 2002.

[23] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Computer vision and pattern recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 519–528. IEEE, 2006.

[24] R. Szeliski. *Computer vision: algorithms and applications*. Springer, 2010.

[25] E. Tola, C. Strecha, and P. Fua. Efficient large-scale multi-view stereo for ultra high-resolution image sets. *Machine Vision and Applications*, 23(5):903–920, 2012.