GRIFFITH UNIVERSITY

Primitive-based 3D Structure Inference from a Single 2D Image for Insect Modeling: Towards an Electronic Field Guide for Insect Identification

Author

Zhang, X, Gao, Y, Caelli, T

Published

2010

Conference Title

11th International Conference on Control, Automation, Robotics and Vision, ICARCV 2010

DOI

10.1109/ICARCV.2010.5707814

Rights statement

© 2010 IEEE. Personal use of this material is permitted. However, permission to reprint/ republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Downloaded from

http://hdl.handle.net/10072/38084

Griffith Research Online

https://research-repository.griffith.edu.au

2010 11th Int. Conf. Control, Automation, Robotics and Vision
Singapore, 7-10th December 2010

Primitive-based 3D Structure Inference from a Single 2D Image for Insect Modeling

Towards an Electronic Field Guide for Insect Identification

Xiaozheng Zhang[†‡], Yongsheng Gao[‡†], Terry Caelli[†]

[†]Biosecurity Group, Queensland Research Lab
National ICT Australia
Brisbane, Australia
{paul.zhang, yongsheng.gao, terry.caelli}@nicta.com.au

[‡]Computer Vision and Image Processing Lab,
Griffith University
Brisbane, Australia
{x.zhang, yongsheng.gao}@griffith.edu.au

***Abstract*—3D insect models are useful to overcome viewing angle variations and self-occlusions in computer-assisted insect taxonomy for electronic field guides. The acquisition of 3D information is, however, unreliable due to the flexibility and small size of the insect bodies. This paper explores how to infer 3D insect models from a single 2D insect image, which will assist both insect description and identification. The 3D structure of the insect body is modeled from two geometric primitives, generalized cylinders and deformable ellipsoids. The primitives are fitted and warped based on both edge and medial axis constraints of the 2D image. Individualized 3D models are then built to approximate the insect structure. The proposed approach results in seemingly useful 3D insect models capable of representing the major morphological characteristics for a variety of insects with different body types. This method could be a helpful assistance for computer-assisted insect taxonomy and insect identification by entomologists and the public.**

***Keywords*—3D model, insect description, structure inference, 3D reconstruction.**

I. INTRODUCTION

Insects are the most diverse animal group with over one million described species. Since most pests are insects, their identification is important in pest control and food biosecurity. They are also of great interests to the public as insects are pervasive in our environment. Due to increasing lack of expertise to insect identification [6], computer-assisted taxonomy (CAT) is desirable for both pest control systems and electronic field guides. It may rely on 1) morphological characteristics such as shape information, 2) molecular signatures such as 16s DNA [2], 3) mass spectrometry, 4) behavioral traits, and 5) sound [5]. The first trait existed for the longest time and remains the most natural and widely used method in taxonomy, although DNA-based method could provide more accurate results. Morpho-taxonomy will continue to be at least the backbone of taxonomic work, especially for an electronic field guide system of insect identification.

With recent advances in computer vision and pattern recognition research, a few attempts have been made based on insects' morphological characteristics towards computer-assisted taxonomy. Weeks et al. [15] applied image-based pattern recognition techniques on insect wings to classify *Ichneumonidae*. This image-based strategy was later transformed into a CAT system known as DAISY (Digital Automated Identification SYstem) [14, 16]. Recently, Mayo and Watson [10] applied automatic image-based feature extraction and machine learning techniques on the identification of live moths and achieved about 85% accuracy on a dataset of 774 individuals over 35 species. These holistic image-based approaches focus on primarily identification of closely related species, because holistic image comparison is often sensitive to image variations caused by species difference. Larios et al. [9] proposed to extract local image features to form concatenated histograms for recognition of deformable stoneflies. The use of local image features provides a better tolerance over body deformation as well as certain kind of image variations due to species differences. Those techniques are 2D image-based and suffer from viewing angle variations and in-depth self-occlusions.

To overcome these limitations, it is desirable to acquire 3D insect models for both descriptions and identifications. An identification system gains great benefits even from a rough 3D insect structure, because it can help to rectify viewing angles and compensate articulated body parts, especially for wings, legs and antennas. Due to self-occlusion and the small body size of insects, traditional 3D scanning cannot work properly. Previous reconstruction methods model 3D structures from shading [18], contour [13] and texture [3, 12], using 2D images.

978-1-4244-7815-6/10/$26.00 ©2010 IEEE

Recently, a few interactive modeling techniques have been proposed to infer 3D object structures from single 2D images or sketches. Gingold et al. [7] developed a user annotation system based on geometric primitives for 3D modeling. From different parts of the bodies, users control and change the parameters of 3D geometric primitives to best fit the input images or sketches. Wu et al. [17] transfer reference normal distributions to target 2D image based on image edges and user inputs. With a reference simple geometry (typically a sphere), users draw the paths on both the object image and sphere image whose normals they think are identical. With pixel-wise interpolations, normals as well as structures from the normal map of the entire object are transferred from the reference object. These innovative ideas and techniques inspired us to develop a 3D structure inference approach for 3D insect modeling which requires only a single insect image and prior knowledge of human beings about which parts of the insect body to be modeled as a certain type of geometric primitives.

To better facilitate insect descriptions and identifications, this paper proposes to infer 3D insect structure from a single 2D image with the help of geometric primitives and automatic/interactive image feature extraction. It operates on real insect images with assistive assignment of meaningful body parts to either a generalized cylinder or an ellipsoid by the user. Then the primitives are deformed based on extracted image features to best fit individual insect on the image. Different body parts are then combined to represent the entire 3D insect model. These reconstructed 3D insect models can then be illuminated, rotated, projected, as well as being deformed to simulate body articulations if appropriate joints are assigned. Compared to traditional modeling techniques such as [13], the assistive geometric constraints are semantic part-based, providing the proposed method higher robustness and accuracy in 3D modeling. The whole process results in good 3D structure inferences, as well as parameterization of 3D insect body useful for further image understanding and pattern recognition processing.

The rest of the paper is organized as follows. Section 2 elaborates the proposed 3D inference approach with an example of modeling *Ptinus exulans*. It first discusses the different connection types of insect body parts. Then 3D inference methods using generalized cylinders and deformable ellipsoids are presented, respectively. Section 3 shows the 3D modeling results from single 2D images of a variety of insects using the proposed 3D inference approach. The conclusions and discussions are given in Section 4.

II. 3D STRUCTURE INFERENCE

All insect images in this paper are taken from Pests and Diseases Image Library (PaDIL) [1]. Figure 1 shows the dorsal view of *Ptinus exulans* with automatic boundary extractions with a few user interactive corrections. Two 3D geometric primitives, generalized cylinders and deformable ellipsoids, are introduced to model different parts of insect body based on the shapes of different body parts. For this insect case, the head and abdomen are modeled as ellipsoids, while thorax, legs and antennae can be properly approximated using generalized cylinders. Before presenting the specific modeling methods using these geometric primitives, the connections of body parts will be first discussed which affect the 3D modeling of the entire insect body.



Figure 1. 2D image of *Ptinus exulans* with extracted boundary (blue) from PaDIL.

A. Connection types of body parts

Different body parts are separated yet inter-connected with different connection types. To facilitate the modeling process, two types of connection are considered and thus processed differently, i.e., 1) normal connection, and 2) co-medial axis connection. Co-medial axis connection shares at least C^0 continuous medial axis curve between adjacent body parts. For instance, as shown in Figure 2, femur (equivalent to thigh), tibia (equivalent to shin) and tarsus (foot) of the right foreleg are co-medial axis connected. Although the boundary curves may change abruptly, the medial axis curve is continuous. For such connected parts, the treatments are different for medial axis curve fitting and boundary-based 3D inference. To fit splines to medial axis pixels, at least C^0 smoothness is enforced. C^1 or C^2 continuity can be enforced, if no big articulation is present as shown in Figure 2. To model generalized cylinder based on medial axis spline and boundary, body parts are then treated with bigger allowances of curvature changes between leg segments. For normal connection type, both medial axis and boundary are almost always discontinuous and are treated as such.

In Figure 1, co-medial axis connections exist between femurs and tibias, between tibias and tarsus, between head and thorax, and between thorax and abdomen. The connections other than them are normal connections, such as the connection between antenna and head. The introduction of connection types follows human perceptions. Only with help of connected head and abdomen, for instance, one can confidently tell the orientation of thorax.

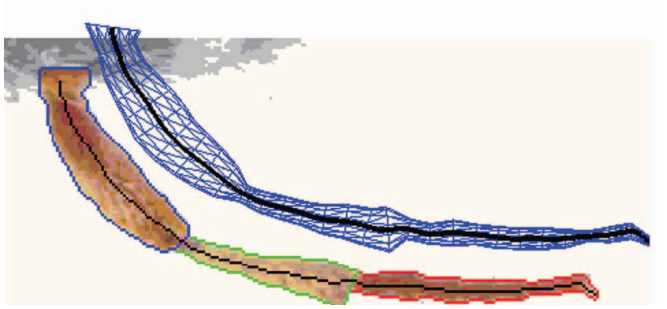


Figure 2. Medial axis is a continuous curve across femur, tibia and tarsus of an insect leg. Blue mesh is the 3D inferred structure from the image, with continuous medial axis and discontinuous 3D surfaces on joints.

B. 3D body part inferring using generalized cylinders

After defining connection types for all body connections, the body parts are separately processed. To model cylindrical body parts such as antennas and legs, circular cross-section generalized cylinders are used. Each cylinder consists of two non-uniform cubic B-splines $S(u) \in \mathbb{R}^3$ and $B(u) \in \mathbb{R}$, $u \in [0,1]$. 3D spline $S(u)$ is approximated from the medial axis points of the body part with interactive in-depth specification, which forms the medial axis curve of the generalized cylinder. Scalar spline $B(u)$ specifies the radii of circular cross-sections for the generalized cylinder which controls the deviation of the 3D surface from its medial axis and is estimated from the distance between the boundary pixels and the pixels on its medial axis.

Given a sequenced set of n medial axis points $\boldsymbol{Q}_k$ $(k = 1, \dots, n)$ extracted from an insect body part, e.g., from the rear left leg as shown in Figure 2, control points $\boldsymbol{P}_i$ $(i = 1, \dots, m)$ are to be estimated with predefined parameters $\bar{u}_k$ corresponding to $\boldsymbol{Q}_k$ and knots u_i corresponding to $\boldsymbol{P}_i$. The chord length and averaging methods are used, respectively, to define $\bar{u}_k$ and u_i [11]. Set $\bar{u}_1 = 0$, then

$$\bar{u}_k = \bar{u}_{k-1} + \frac{\|\vec{Q}_k - \vec{Q}_{k-1}\|}{\sum_{j=2}^{n} \|\vec{Q}_j - \vec{Q}_{j-1}\|}, \quad k = 2, \dots, n. \tag{1}$$

For knots, set $u_1 = \cdots = u_4 = 0$ and $u_{n+1} = \cdots = u_{n+4} = 1$, and the knot u_i $(i = 5, \dots, n)$ in between are calculated by averaging the adjacent 4 parameters $\bar{u}_{i-4}, \dots, \bar{u}_{i-1}$ as

$$u_i = \frac{1}{4} \sum_{j=i-4}^{i-1} \bar{u}_j, \quad i = 5, \dots, n. \tag{2}$$

Because a dense spline is often affected by noisy medial axis data, sparse knots $\tilde{u}_j$ with interval l_s are chosen to ensure a smooth spline approximation. $m = \frac{n-2}{l_s} + 6$ control points $\{\boldsymbol{P}(\tilde{u}_j)\}$ are then calculated by minimizing the objective function which measures the different of sampled points $\boldsymbol{Q}_k$ and its counterparts on spline $S(u)$, as

$$g\left(\boldsymbol{P}(\tilde{u}_j)\right) = \sum_{k=1}^{n} [\boldsymbol{Q}_k - \boldsymbol{S}(\bar{u}_k)]. \tag{3}$$

The fitted medial axis B-spline is shown in Figure 2 (black). By evaluating $\boldsymbol{S}(u)$, the distances $d_l(u)$ and $d_r(u)$ from boundary points of two sides are recorded and the average distance is used as the sample data for $B(u)$. Similar to $\boldsymbol{S}(u)$, but with a shorter knot interval $\mathrm{l_b}$, $B(u)$ is estimated by approximating the average distances and used as the radii of the cross-sections of the generalized cylinder (blue mesh in Figure 2). By this means, different connection types are treated differently in the forms of two separate splines.

C. 3D body part inferring using deformable ellipsoids

For those body parts to be modeled as ellipsoids (e.g., head and abdomen in this case), the boundary pixels are extracted and fitted using an ellipse by minimizing

$$f(\boldsymbol{a}) = \sum_{i=1}^{N} (\boldsymbol{x}_i \cdot \boldsymbol{a}), \tag{4}$$

subject to $b^2 - 4ac < 0$, where $\boldsymbol{x} = [x^2, xy, y^2, x, y, 1]$, (x, y) is a boundary point, and $\boldsymbol{a} = [a, b, c, d, e, f]$ is the ellipse defining vector [8]. Ellipse coefficients are then computed, such as center, semi-axes lengths (a', b') and major axis rotation angle. A hemiellipsoid is constructed by enforcing circular cross-sections perpendicular to the major axis, as

$$\begin{cases} x' = a' \cos\theta \\ y' = b' \sin\theta \cos\phi \\ z' = b' \sin\theta \sin\phi \end{cases} \tag{5}$$

where $\theta, \phi \in [0, \pi]$ are elliptic and circular parameters, respectively.

For the ellipse $\boldsymbol{X_s} = \{x_i', y_i'\}$ of xy plane, we establish a corresponding point set on the boundary, which corresponds the closest points on the boundary to the points on the ellipse, as

$$\boldsymbol{X_t} = \{x_i, y_i | \min \|(x_i, y_i) - (x_i', y_i')\|\}. \tag{6}$$

A thin-plate spline warping [4] is performed on the inferred hemiellipsoid to fit boundary points by specifying $\boldsymbol{X_s}$ and $\boldsymbol{X_t}$ as source and target control points, respectively. No smooth term is needed because 1) a perfect fitting is desirable, and 2) minimal distances between source and target control points are used.

Figures 3-5 shows the process of ellipsoid modeling and 3D structure transfer using thin-plate spline warping on the abdomen of *Ptinus exulans* (Figure 1). Based on the extracted boundary pixels, an ellipse is first fitted using (4), as shown in Figure 3. Then Figure 4 shows 3D in-depth information inferred by enforcing circular cross-sections along medial axis spline. A "perfect" hemiellipsoid (blue) is built, which does not necessarily follow the boundary pixels. Using sampled points on the ellipse and their closest boundary points as control points of source and target respectively, a thin-plate spline warp

is performed to transfer the 3D model to best fit the input image (boundary), as shown in Figure 5. The meshes before and after structure transfer are shown in blue and red, respectively. During this process, only x and y coordinates of the vertices change while z coordinates are fixed. So compared to original mesh, new mesh follows boundary points more faithfully.

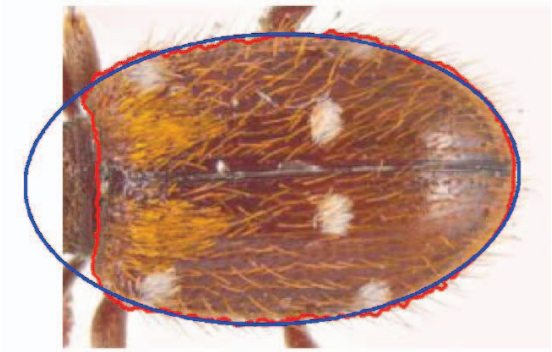


Figure 3. Image-based ellipse fitting for insect abdomen.

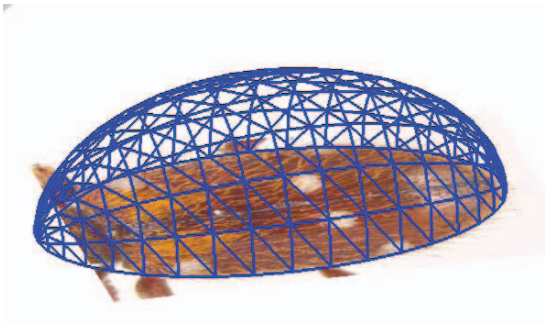


Figure 4. 3D depth inference of insect abdomen using a hemiellipsoid.

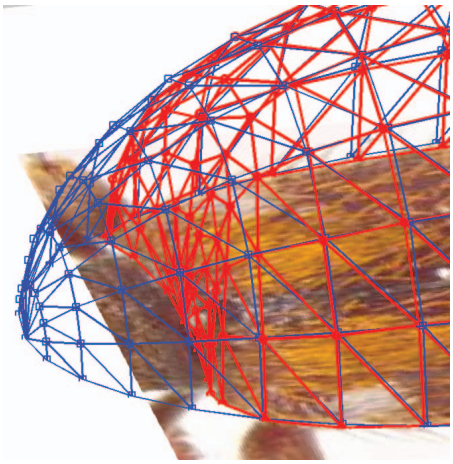


Figure 5. Boundary-based structure transfer using thin-plate spline warping.

Different body parts are then modeled from either generalized cylinders or deformed ellipsoids. A typical body configuration includes a head, a thorax, an abdomen, 6 legs, and 2 antennas. These body parts are integrated into an overall 3D model of the insect from a single 2D insect image (Figure 1) as shown in Figure 6.

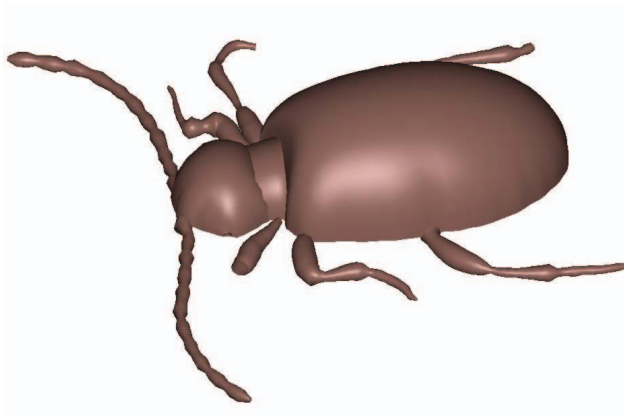


Figure 6. 3D inference result of *Ptinus exulans* from a single 2D image (Figure 1).

III. 3D Insect Inference Results

We have applied our proposed 3D inference approach to model a variety of insects from a single 2D dorsal image. All these images are from PaDIL database which is publicly available on the internet and these images serve for identification purposes for both general public and entomologists. These are shown in Figure 7 and Figure 8, and the species are *Niptus hololeucus* (golden spider beetle) and *Chrysophtharta variicollis* (eucalyptus leaf beetles), respectively. User interactions include boundary corrections and choosing appropriate geometric primitives for different parts of the insect bodies. The 3D structures are then automatically inferred from the single 2D images and rendered from different viewing angles.

From the results shown in Section 2 and 3, the choice of using geometric primitives provides the 3D modeling approach additional and appropriate constraints on the inferred 3D structures from a single 2D image. For body parts different from perfect geometric primitives such as the abdomen of *Ptinus exulans*, the proposed approach deforms the geometric primitive to fit the individual insect structure. The generalized cylinders can approximate long antennas and legs of *Ptinus exulans* and *Niptus hololeucus*, as well as the short legs of *Chrysophtharta*. We haven't tried on insects with wings which request us in our future work to consider a new geometric primitive to approximate wing-shaped body parts. However, current approach will be adequate to model a variety of wingless insects, including beetles, ants and termites, which consist of a large number of biologically high impact species. These results show that good choices of geometric primitives with reasonable constraints, such as smoothness and restricted cross-section shapes, provide visually plausible results to an ill-posed problem. Image-based 3D structure transfer helps to

deform primitives to fit to individual insect image, thus improves the modeling quality.



Figure 7. 3D inference result (bottom) of *Niptus hololeucus* from a single 2D image (top).

IV. Conclusions

In this paper, a 3D structure inference approach has been proposed for 3D insect modeling from single 2D pictures towards an electronic field guide of insect description and identification. It incorporates primitive-based modeling and feature-based structure transfers to achieve regularized individual 3D structures from 2D insect images. Dependent on its structural characteristics, each part of insect body is modeled as either a generalized cylinder or a deformable ellipsoid, which serves as a regularization term for this ill-posed problem. For generalized cylinder modeling, both the medial axis and the width of the cylinder are modeled using separate parametric splines, respectively, taking into considerations the different connection types of body parts. For ellipsoid-like body part such as head and abdomen, the model is further warped based on the difference between ellipsoid silhouette and image boundary curve, which deforms the model to fit the individual input image. The results show our proposed method produced plausible and parametric 3D insect models, which could be used to assist computer-aided insect identification for plant biosecurity and electronic field guides.

The future work is to incorporate pixel intensity information into this 3D modeling approach to further improve the accuracy of the model as well as provide reflectivity properties of insect body surfaces.

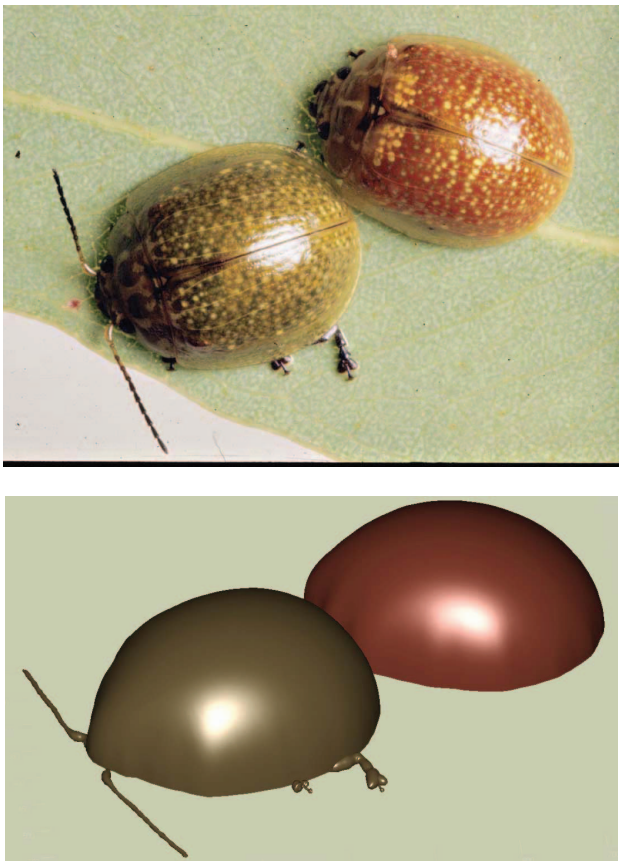


Figure 8. 3D inference result (bottom) of *Chrysophtharta variicollis* from a single 2D image (top).

References

[1] "PaDIL - Pests and Diseases Image Library," in http://www.padil.gov.au/.

[2] S. L. Ball and K. F. Armstrong, "Rapid, one-step DNA extraction for insect pest identification by using DNA barcodes," *Journal of Economic Entomology*, vol. 101, no. 2, pp. 523-532, 2008.

[3] D. Blostein and N. Ahuja, "Shape from texture: integrating texture-element extraction and surface estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 12, pp. 1233-1251, 1989.

[4] F. L. Bookstein, "Principal warps: thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567-585, 1989.

[5] K. Gaston and M. O'neill, "Automatic species identification: why not?," *Philosophical Transactions of the Royal Society B-Biological Sciences*, vol. 359, pp. 655-667, 2004.

[6] K. J. Gaston and M. A. O'Neill, "Automated species identification: why not?," *Phil. Trans. Roy. Soc. B.*, vol. 359, no. 1444, pp. 655-667, 2004.

[7] Y. Gingold, T. Igarashi, and D. Zorin, "Structured annotations for 2D-to-3D modeling," *ACM Trans. Graph.*, vol. 28, no. 5, 2009.
[8] R. Halir and J. Flusser, "Numerically stable direct least squares fitting of ellipses," 1998.
[9] N. Larios, H. L. Deng, W. Zhang, M. Sarpola, J. Yuen, R. Paasch, A. Moldenke, D. A. Lytle, S. R. Correa, E. N. Mortensen, L. G. Shapiro, and T. G. Dietterich, "Automated insect identification through concatenated histograms of local appearance features: feature vector generation and region detection for deformable objects," *Machine Vision and Applications*, vol. 19, no. 2, pp. 105-123, 2008.
[10] M. Mayo and A. T. Watson, "Automatic species identification of live moths," *Knowledge-Based Systems*, vol. 20, no. 2, pp. 195-202, 2007.
[11] L. Piegl and W. Tiller, *The NURBS Book*, 2nd ed. Springer, 1997.
[12] D. Samaras and D. Metaxas, "Incorporating illumination constraints in deformable models for shape from shading and light direction estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 247-264, 2003.
[13] F. Ulupinar and R. Nevatia, "Shape from contour: straight homogeneous generalized cylinders and constant cross section generalized cylinders," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 2, pp. 120-135, 1995.
[14] P. J. D. Weeks, O. N. M. A., K. J. Gaston, and I. D. Gauld, "Development of an automated insect identification system," *Journal of Applied Entomology*, vol. 1, pp. 1-8, 1997.
[15] P. J. D. Weeks, I. D. Gauld, K. J. Gaston, and M. A. O'Neill, "Automating the identification of insects: a new solution to an old problem," *Bull. Ent. Res.*, vol. 87, pp. 203-211, 1997.
[16] P. J. D. Weeks, M. A. O'Neill, K. J. Gaston, and I. D. Gauld, "Automating insect identification: exploring the limitations of a prototype system," *Journal of Applied Entomology-Zeitschrift Fur Angewandte Entomologie*, vol. 123, no. 1, pp. 1-8, 1999.
[17] T.-P. Wu, C.-K. Tang, M. S. Brown, and H.-Y. Shum, "ShapePalettes: interactive normal transfer via sketching," *ACM Trans. Graph.*, vol. 26, no. 3, 2007.
[18] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah, "Shape-from-shading: a survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 8, pp. 690-706, 1999.