MAELab: a framework to automatize landmark estimation

van Linh Le, Marie Beurton-Aimar, Adrien Krähenbühl, Nicolas Parisey

MAELab: a framework to automatize landmark estimation

LE Van Linh
LaBRI-CNRS 5800
ITDLU, Dalat Univ-V
linhlv@dlu.edu.vn/
van-linh.le@labri.fr

BEURTON-AIMAR Marie
LaBRI-CNRS 5800
Bordeaux University
33400 Talence-F
beurton@labri.fr

KRAHENBUHL Adrien
LaBRI-CNRS 5800
Bordeaux University
33400 Talence-F
adrien.krahenbuhl@labri.fr

PARISEY Nicolas
IGEPP
INRA 1349
35653 Le Rheu-F
nparisey@rennes.inra.fr

Abstract

In biology, the morphometric analysis is widely used to analyze the inter-organisms variations. It allows to classify and to determine the evolution of an organism's family. The morphometric methods consider features such as shape, structure, color, or size of the studied objects. In previous works [8], we have analyzed beetle mandibles by using the centroid as feature, in order to classify the beetles. We have shown that the Probabilistic Hough Transform (PHT) is an efficient unsupervised method to compute the centroid. This paper proposes a new approach to precisely estimate the landmark geometry, points of interest defined by biologists on the mandible contours. In order to automatically register the landmarks on different mandibles, we defined patches around manual landmarks of the reference image. Each patch is described by computing its SIFT descriptor. Considering a query image, we apply a registration step performed by an Iterative Principal Component Analysis which identify the rotation and translation parameters. Then, the patches in the query image are identified and the SIFT descriptors computed. The biologists have collected 293 beetles to provide two sets of mandible images separated into left and right side. The experiments show that, depending on the position of the landmarks on the mandible contour, the performance can go up to 98% of good detection. The complete workflow is implemented in the MAELab framework, freely available as library on GitHub.

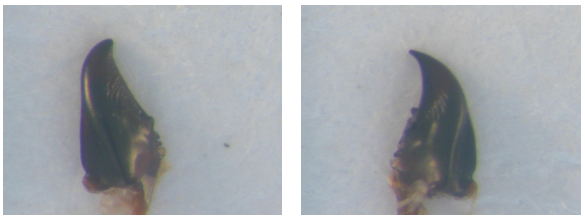
Keywords

Morphology, image registration, SIFT descriptor, beetle, mandible.

1 INTRODUCTION

Phenotype of beetle species is characterized by information like age, sex, morphological criteria or environmental parameters. Biologists are used to proceed to manual measurements in case of analysis at macro level as for example tissues or animal members [6] [3]. They can directly measure the geometrical characteristics of elements on the body of the animal: length, width, diameter, angles, etc. Another way to obtain morphological measures is to take pictures of the members and to apply image processing algorithms. In order to evaluate a population of beetles from Brittany lands, a collection of 293 beetles has been established. For each beetle, biologists took images of the left and right mandibles (see Fig. 1) and a set of landmarks has been manually determined by experts. A morphometric landmark is a specific point, directly linked to the animal anatomy.

The landmarks are used in many domains, not only in biological studies. It is an application field of image processing [9] that appears in works of computer vision, mainly in face recognition [15] but also in human orthodontic [5] or morphometric analysis [1].



(a) Left mandible (b) Right mandible

Figure 1: The two mandibles of a beetle captured by biologists.

In this paper, we focus on the automatic identification of landmarks in 2D mandible images. The proposed method consists of three main steps: a segmentation of the mandible based on the Canny algorithm, an Iterative Principal Component Analysis to register a query image on a model image, and finally a landmark estimation on the query image by comparison of SIFT descriptors. Section 2 presents the complete workflow then section 3 details the experiments and analyzes the results.

2 METHOD

The addressed problem is the automatic detection of morphologic landmarks on mandible pictures to replace the manual operation made by an expert operator. We detail hereafter a workflow (resumed Fig. 2) including (1) the segmentation of a query image, (2) a registration step on a model image then (3) the detection of landmark positions. It is worth to note that all pictures have been taken in the same conditions with the same camera at the same resolution. Moreover, the model image is randomly chosen from the set of all images.

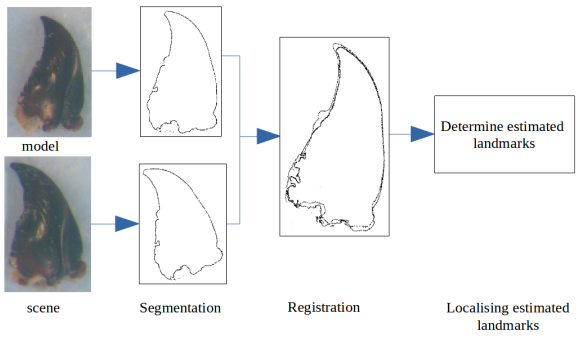


Figure 2: Overview of the proposed method.

2.1 Image segmentation

The segmentation step is the first task of a large number of image processing chains. A contour-based algorithm, the Canny algorithm [4], has been chosen to extract the contour belonging to the shape of the mandible. To use this method, two threshold values have to be set. As it is mentioned in [14], determine the right thresholds could be difficult. The mandatory *threshold value* is determined by analyzing the image histogram (see [8] for details). Most often authors define these thresholds as a lower and an upper one with a usual ratio of $T_{lower} = (1/2) \times T_{upper}$. In order to consider a larger range of values, we defined $T_{lower} = 1/3 \times T_{upper}$. Note that to optimize the computing time, the direction of the gradient of each pixel belonging to the mandible contour is also computed during the Canny algorithm but will be used later (see Sec. 2.3). To obtain the segmentation of the mandible, the contours obtained with Canny are discriminated to only keep the mandible contours. As shown in Fig. 3, the Canny algorithm generates some contours which do not belong to the mandible shape. A simple algorithm parses the contour image to suppress the edges inside the biggest contour.

2.2 Image registration

As previously mentioned, all images have been captured at the same scale. However, the mandible size can vary from a beetle to another one, as their orientation and position can differ from a picture to another

(a) Contours identified by the Canny algorithm.

(b) Contour selected by post-processing.

Figure 3: Detection of the mandible contours.

one. This point is taken into account in this registration step from a query image (the scene) to a reference image (the model).

We have chosen to apply a method based on the Principal Component Analysis (PCA) [12, 13] to determine the rotation and translation parameter between the two images. In input, we consider the two lists of contour points defined from the segmentation of the two images. Firstly, the centroid and the principal axis of each image are computed. The centroid corresponds to the mean point of all contour points. The principal axis is the line connecting the centroid to a contour point, determined with algorithm 1.

Algorithm 1: Algorithm to find the principal axis of a list of contour points

Input : Centroid c, list of contour points l
Output: Principal axis a

```
for all points p_i in l do
    for all points p_j in l do
        if p_i ≠ p_j then
            Compute the orthogonal distance d_ij
            between the line (c, p_i) and p_j.
        end
    end
    Compute d_mean as the average distance of all d_ij
    distances.
    if d_mean is minimal then
        p_min = p_i;
    end
end
The principal axis is: a = (c, p_min).
```

The translation between the scene image and the model image is computed from the distance between their centroids. The rotation is computed from the angle between the principal axes of these two images. Translation then rotation operations are applied to register the

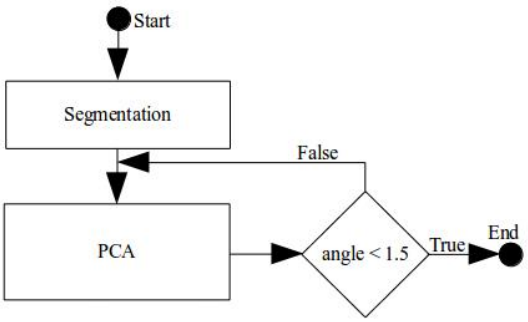


Figure 4: Workflow of the PCAI, allowing to refine the rotation angle between the scene and the model.

scene on the model. However, the translation and rotation can be imprecise when the result of the segmentation contains noise. In order to prevent these cases, we remarked and used the image specificity that implies that the tip of the mandible is less noisy than its base. So, we have sorted the contour points according to their y-value to build the subset containing the half upper part of contour points. Then, we have enhanced the PCA by an Iterative process on this subset until stabilization (PCAI). At each iteration, the rotation value is refined. The procedure stop when the new computing angle is lower than 1.5 degrees (see Fig. 4).

The Fig. 5 shows an example of the successive results obtained with PCAI. The red contour belongs to the model, the black one corresponds to the scene contour registered according to the angle find after one iteration, and the blue contour is the final contour obtained at the end of the PCAI process.

Figure 5: Iterations of the registration step between the model contour (in red) and the contours of the scene image.

2.3 Landmark detection with SIFT

The last step of the workflow consists of estimating the position of the landmark in scene image from the manual ones of the model. We relied on the SIFT method [10] that has been used in a lot of computer vision methods to identify points of interests. We modified some aspects of the initial method to defined a process specialized to the landmark identification. In order to reduce the computing time and the possible errors of location, we do not consider all points of the image but only the area around each landmark on the model. Firstly, the patch around each landmark of the model is computed and extracted at the same position in the scene image. Then, the SIFT descriptor is computed. The orientation and the gradient magnitude are calculated for each pixel by using the gradient values computed during contour detection (see Sec. 2.1) by applying the classical equations 1:

$$\begin{aligned} m_I(x,y) &= \sqrt{v_x^2 + v_y^2} \\ \theta_I(x,y) &= tan^{-1}(v_y/v_x) \end{aligned} \tag{1}$$

Where:

- $v_x = I(x+1,y) - I(x-1,y)$
- $v_y = I(x,y+1) - I(x,y-1)$
- $I(x,y)$ is the intensity of I at position (x,y),
- $m_I(x,y)$ is the gradient magnitude of in I at position (x,y),
- $\theta_I(x,y)$ is the orientation in I at position (x,y).

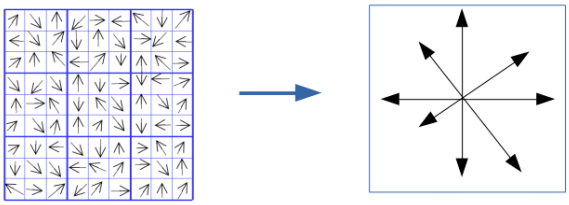


Figure 6: The SIFT descriptor of a patch. In the right figure, the arrow length corresponds to the gradient value.

For each patch, the SIFT descriptor is the histogram of the sum of gradients of each considered direction. As usually, eight direction classes are considered: $[0^o - 45^o[$, $[46^o - 90^o[$, $[91^o - 135^o[$, $[136^o - 180^o[$, $[181^o - 225^o[$, $[226^o - 270^o[$, $[271^o - 315^o[$, $[316^o - 360^o[$). The feature vector is normalized to reduce the effects of illumination changes.

The Fig. 6 shows a patch of 9×9 pixels centered in each landmark on the model. The size of 9×9 has been retained after several tests where patch sizes 18×18, 36×36 and 54×54 have given unsatisfactory results. From the histogram, we obtain the local gradient value for each direction.

The comparison between two SIFT descriptors is done by using the $L2$-distance with the following equation (2):

$$L(D1,D2) = \sum_{i=0}^{n} \sqrt{(D1_i - D2_i)^2} \tag{2}$$

Where:

- n is the number of directions
- $D1$ and $D2$ are two descriptors of size n,
- $D1_i$ and $D2_i$ are the i^{th} descriptor value.

The Fig. 7 illustrates how we have applied the SIFT method into our workflow. IN order to detect the scene landmarks, the patches P_m of the model and P_s of the scene are initialized with the size of P_m smaller than the size of P_s. After experiments, we have kept 36×36 pixels as the size of P_s. For each pixel in the patch P_s, a sub-patch P'_s is extracted with the same size than P_m. When P'_s have a part outside P_s, the outside pixels are also considered. Then, the distance $L(P_m, P'_s)$ is computed using equation (2). The estimated position of the landmark corresponds to the position of the sub-patch P'_s with the smallest distance L to P_m. Finally, the position of the estimated landmarks are positioned in the original scene image by applying the reverse operations of rotation and translation computed in registration step (see Sec. 2.2).

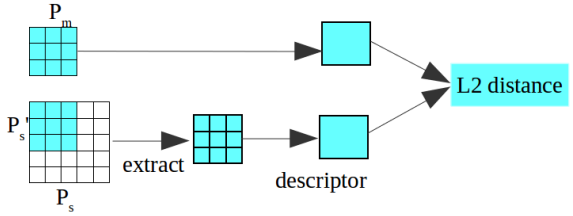


Figure 7: Comparison of descriptors between the patch P_m of the model image and the patches P'_s of the scene image.

3 EXPERIMENTS AND RESULT

The complete method is implemented in the framework MAELab[1]. The left and the right mandibles of the beetles has been analyzed separately. After verifying the quality of the image, it remains 290 usable images of right mandibles and 286 images of left mandibles. The removed images include the images without mandible or with broken mandibles. In all valid images, a set of manual landmarks is indicated by biologists: 18 for right mandibles, 16 for left mandibles, which constitutes our ground truth.

We have run the full workflow on all the usable images. The results have shown differences in algorithm accuracy: estimated landmarks are well positioned on some scene images but not satisfying on others. As we mentioned before, mandibles images can exhibit different sizes because beetles have also different sizes of mandible. We detected that our method is sensible to this parameter. To improve the results, we have inserted a pre-processing step to estimate the scale of a scene image and the model before the computing of the SIFT descriptors. The bounding boxes of the mandible of the model image and the scene image are computed and the scales in the x- and y-directions are determined by the ratio between the corresponding sides of the bounding boxes. Then, the scene contours are rescaled to fit the model contours.

[1] MAELab is a free software written in C++. It can be directly and freely obtained by request at the authors.

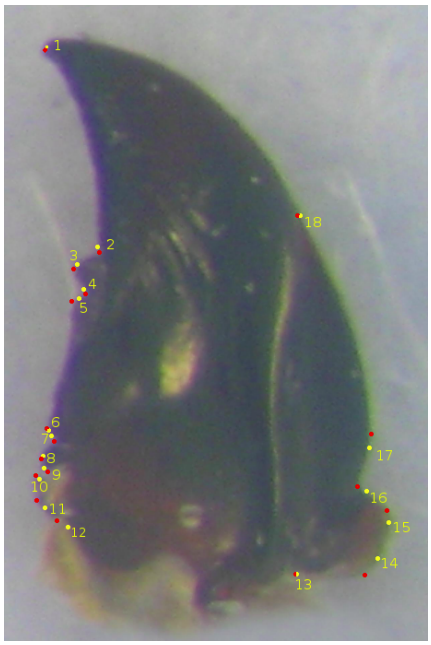


Figure 8: The manual (in red) and estimated (in yellow) landmarks on a right mandible.

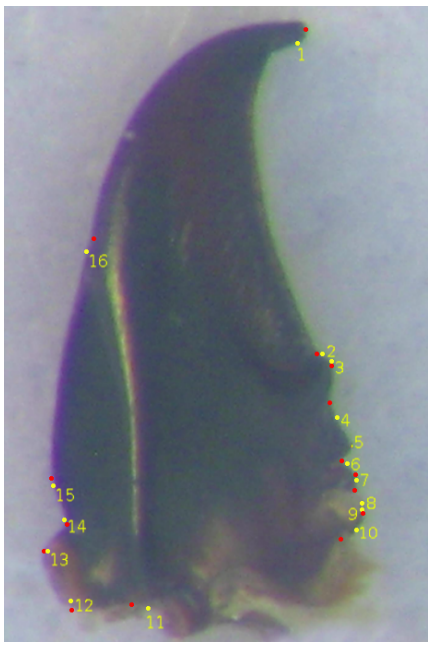
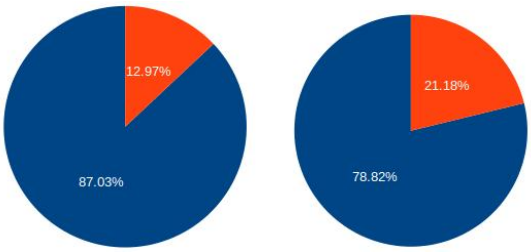


Figure 9: The manual (in red) and estimated (in yellow) landmarks of a left mandible.

Figs. 8 and 9 show the final results for a right and a left mandible with the manual and estimated landmarks. The estimated landmarks are quite near with the manual landmarks, as it is shown in the following statistical evaluation.

The statistics have been computed for all landmarks of the scene images. We have compared the positions between the manual and estimated landmarks by accepting an error from 1% to 2% of the bounding box's size. According to this way, a global statistic compares all pairs of corresponding landmarks on all images as presented in Fig. 10. It shows the global results with a score of well-positioned landmarks equal to **87.03**% for right mandibles and **78.82**% for left mandibles.



(a) Set of right mandibles. (b) Set of left mandibles.

Figure 10: The mean proportion of well and bad landmark locations of the two sets of left and right mandibles.

Besides the global results, we are also interested by the accuracy of the individual positions of the estimated landmarks. We have computed the distance between the manual landmarks and their corresponding estimated landmarks in order to examine the proportion of well-positioned landmarks. The Fig. 11 and 12 show the proportion of well-estimated landmarks for each landmark of the model. With 18 landmarks of right mandible, the position of the 1^{st} estimated landmarks is very accurate with **98.62**%. The lowest proportion is **74,48**% for the 14^{th} landmark. The remaining landmarks are also estimated with an accuracy greater than 75%. For left mandibles, the highest and lowest success rates are **93,01**% for the 1^{st} landmark and **60,14**% for the 16^{th} landmark. The statistic is done on each estimated landmark of all the images with a standard deviation error [2]. As we can see in Fig. 3, the noise of the contour part located at the base of a mandible is higher than the noise located at the tip of the mandible. This explains why the correct proportion on 11^{th} and 12^{th} landmarks of the left mandible and 13^{th} and 14^{th} landmarks of the right one are less accurate than other landmarks. Moreover, when we reconsider the datasets, the left mandible images have bigger scale values than the right mandible images. This could explain that the success rate of the right mandibles is always greater than this one of the left in all experiments.

In a previous work, we have tried to apply a set of procedures coming from an article of Palaniswamy [11] who tried to find automatically a specific point of interest into a Drosophila wing. We have succeeded to fix the centroid of the mandibles by using these procedures (mainly based on the computation of a Probabilistic Hough Transform accumulator). But this way has not been enough efficient to set precisely the landmarks. D. Houle et al [6] have more recently described a method to estimate automatically the landmarks on Drosophila wings (with 12 landmarks). This method is mainly based on the use of a curve analysis (with splines) belonging to the wing shape. The method has been evaluated on 535 wing images. The average proportion of all 12 points is 82%. They have been able to improve their results by suppressing the least accuracy point (47% of right results) that leads to a better parameter fitting. Y. Ke et al. [7] have proposed to combine SIFT descriptor with PCA analysis to characterize images belonging to a Graffiti dataset. They also obtained good performances close to 95% of correct results. The results presented in this article can be considered as in the same order of correctness than these works, but it concerns a problem, precise fixing of a lot of landmarks, more difficult to solve. One can note that this chain of treatments dedicated to the estimation of landmarks on 2D images of mandibles is from now, user-friendly available.

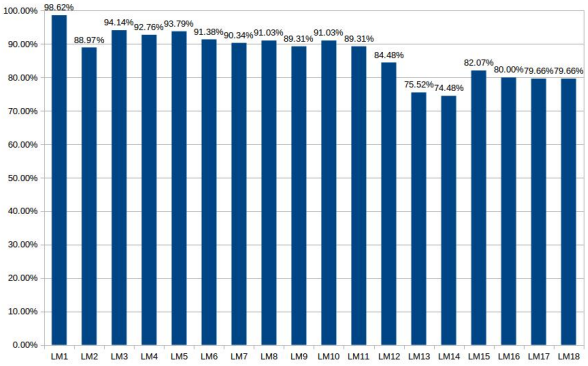


Figure 11: The proportion of well estimated landmarks of right mandibles.

4 CONCLUSION

The morphometric analysis is a powerful tool to analyze and to classify species. In this paper, we have designed a method to segment the beetle mandibles and to automatically locate landmarks which have been determined manually on a model image, by biologists. Each mandible has been segmented by using the Canny algorithm before to be registered using PCAI to align the images. The estimation step of the landmark position uses the SIFT descriptor to find the best matching position. The results show that the method succeeds in locating the landmarks for all images. The accuracy of the method is sufficient to be proposed to biologists as

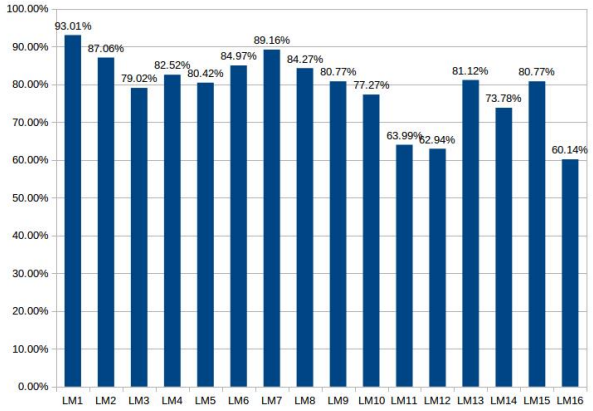


Figure 12: The proportion of well estimated landmarks of left mandibles.

an alternative to the manual measures. Moreover, considering the previous work in [8], this method reduces the drastically the number of outlier landmarks and the MAELab implementation also reduce the global computing times and memory cost. From now, the next stage consists of improving the registration step in order to increase the matching step accuracy and completely remove manual interventions. For example, we could investigate deep learning methods, more precisely Convolutional Neural Networks computing, which has risen up in image processing recently. Biologists are interested in the large-scale analysis of their species collections, automatic classification is one of the bottlenecks to solve towards a better integration of informatic procedures in their current way to work.

5 REFERENCES

[1] José Maria Becerra and Antonio G Valdecasas. Landmark superimposition for taxonomic identification. *Biological Journal of the Linnean Society,*, 81:page 267–274, 2004.

[2] J Martin Bland and Douglas G Altman. Statistics notes: measurement error. *Bmj*, 313(7059):744, 1996.

[3] Paul A Bromiley, Anja C Schunke, Hossein Ragheb, Neil A Thacker, and Diethard Tautz. Semi-automatic landmark point annotation for geometric morphometrics. *Frontiers in Zoology*, 11(1):61, 2014.

[4] John Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698, 1986.

[5] Leila Favaedi and Maria Petrou. Cephalometric landmarks identification using probabilistic relaxation. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 4391–4394. IEEE, 2010.

[6] David Houle, Jason Mezey, Paul Galpern, and Ashley Carter. Automated measurement of drosophila wings. *BMC evolutionary biology*, 3(1):25, 2003.

[7] Yan Ke and Rahul Sukthankar. Pca-sift: A more distinctive representation for local image descriptors. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–II. IEEE, 2004.

[8] L Lê Vãnh, M Beurton-Aimar, JP Salmon, A Marie, and N Parisey. Estimating landmarks on 2d images of beetle mandibles. *WSCG*, 2016.

[9] Yunpeng Li, David. J Crandall, and Daniel P. Huttenlocher. Landmark classification in large-scale image collections. pages 1957–1964. Kyoto, Japan, 29 sept-2 oct 2009.

[10] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[11] Sasirekha Palaniswamy, Neil A Thacker, and Christian Peter Klingenberg. Automatic identification of landmarks in digital images. *IET Computer Vision*, 4(4):247–260, 2010.

[12] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6):559–572, 1901.

[13] Jonathon Shlens. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.

[14] Jun Zeng and Dehua Li. An adaptive canny edge detector using histogram concavity analysis. *International Journal of Digital Content Technology and Its Applications*, 5(6):172–181, 2011.

[15] Zhanpeng Zhang, Ping Luo, Chen Change Loy, and Xiaoou Tang. Facial landmark detection by deep multi-task learning. In *European Conference on Computer Vision*, pages 94–108. Springer, 2014.