

MAELab: a framework to automatize landmark estimation

LE Van Linh	BEURTON-AIMAR Marie	KRAHENBUHL Adrien	PARISEY Nicolas
LaBRI-CNRS 5800 ITDLU, Dalat Univ-V linhvlv@dlu.edu.vn/ van-linh.le@labri.fr	LaBRI-CNRS 5800 Bordeaux University 33400 Talence-F beurton@labri.fr	LaBRI-CNRS 5800 Bordeaux University 33400 Talence-F adrien.krahenbuhl@labri.fr	IGEPP INRA 1349 35653 Le Rheu-F nparisey@rennes.inra.fr

Abstract

In biology, morphometric analysis is widely used to detect the inter-organisms variations. It allows to classify and to determine the evolution of an organism family. The morphometric methods may be divided into groups considering the features used such as shape, structure, color, pattern or size of the object. In a previous work [8], we have shown that the Probabilistic Hough Transform (PHT) method is efficient to compute automatically the centroid position of the beetle mandibles which is used to compare different classes of beetles. In this new work, we address the problem to consider more precisely the position or the geometry of landmarks (points of interest) provided by biologists on a source (i.e model) image, in order to register them automatically on different scene images. Patches around manual landmarks on the model are extracted and the SIFT descriptors of the pixels belonging to the patches are computed. After a registration step between the model and the considered scene, the corresponding patches on the scene are identified and the SIFT descriptors are calculated. To obtain a good registration between the scene and the model, we have defined an Iterative Principal Component Analysis step which sets the rotation and translation parameter values. 290 beetles have been collected and the set of the mandibles images (left and right) has been analysed. The obtained results show that depending on the position of the landmarks on the shape, a very good result can be obtained, i.e more than 98% of the landmarks are rightly placed, or a lower one, i.e 63% for the hardest to fix. The complete workflow is implemented in the MAELab framework, freely available as library on a GitHub website.

Keywords

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

1 INTRODUCTION

Phenotype of beetle species are characterized by informations 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, ... 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 built. For each beetle, biologists took images of the left and right mandibles (see Fig. 1) and a set of landmarks has been manually lo-

cated. The morphometric landmarks are precise points defined by biologists. Landmarks are widely used in many domains, not only in biological studies. It is a large field in image processing [9] and appears in most topics of computer vision such as face detection [13], human orthodontic[5], morphometric analysis [1]



Figure 1: Sample of pictures of beetle mandibles took by biologists.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

In this paper, we focus on a chain of algorithms to automatically identify landmarks in 2D images. The method mainly includes three steps: firstly, a segmentation of the mandible shape by using the Canny algorithm, then an Iterative Principal Component Analysis to register a query image on a model image, and finally a landmark estimation using the SIFT descriptor.

The section 2 presents the complete workflow then the section 3 details the experiments and exposes the results.

2 METHOD

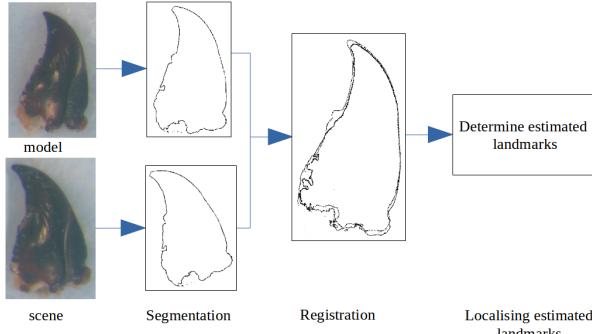


Figure 2: Overview of the proposed method

The addressed problem is the automatic detection of landmarks on mandible photographies to replace the manual operation made by an operator. We propose hereafter a workflow including (1) the segmentation of each image, (2) a registration step on a model image then (3) the detection of landmark positions (Figure 2). It is worth to note that all pictures of mandibles have been taken in same conditions with same camera and same resolution, and that the model image has been chosen randomly from the set of images.

2.1 Image segmentation

The segmentation step is the first major task of a large number of image processing chains. A contour-based algorithm, the Canny algorithm [4], has been chosen to determine 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 [7], determine the right thresholds could be difficult. The mandatory *threshold value* used by Canny algorithm is determined by analyzing the image histogram (see [8] for detail). 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 prefer to set $T_{lower} = 1/3 \times T_{upper}$. One can note that to optimize the computing time, the direction of the gradient of each pixel belonging to the mandible contour is computed during the Canny algorithm and will be used later. 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. With a simple algorithm, the contour image is parsed to suppress the edges inside the bigger contour.

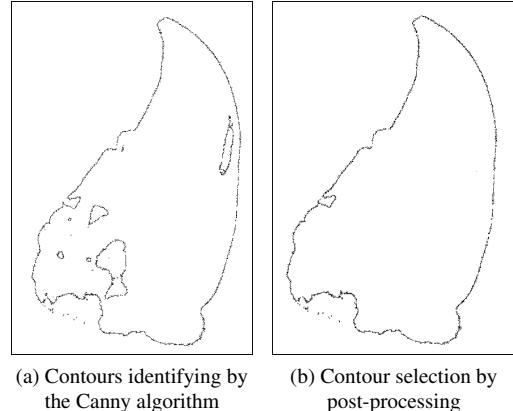


Figure 3: Step of the mandible contour detection.

2.2 Image registration

As it is mentioned previously, all images have been captured with the same scale but the size of mandibles can vary from a beetle to another one as their orientation and position can differ from a photography to another one. This point will be taken into account at this step which concerns the registration of an image (the scene) on a reference image (the model).

We have chosen to apply a method based on the Principal Component Analysis (PCA) [11], [12] to determine the rotation and translation parameter values between the two images. As input values, we use the lists of contour points defined at the segmentation step. Firstly, the centroid point and principal axis of each image are defined. The centroid point coordinates are computed like the mean coordinate of all contour points. The principal axis is a line connecting the centroid point to a contour point, determined as detailed in the algorithm 1.

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

Input : Centroid point c , list of contour points l
Output: The principal axis a

```

1 for all points  $p_i$  in  $l$  do
2     for all points  $p_j$  in  $l$  do
3         if  $p_i \neq p_j$  then
4             Compute the perpendicular distance  $d_{ij}$ 
5                 between line  $(c, p_i)$  and  $p_j$ .
6         end
7     end
8     Compute the average distance ( $d_{mean}$ ) of all
9     distances  $d_{ij}$ ;
10    if  $d_{mean}$  is minimal then
11        |  $p_{min} = p_i$ ;
12    end
13 end
14 The principal axis is:  $a = (c, p_{min})$ .

```

The translation between the scene and the model image is calculated like the distance between their centroid points. The rotation angle is the angle between the principal axes of these two images. Translation and rotation operations are applied to register the scene on the model. However, in some cases, the translation and rotation between two images are not enough precise because the result of the segmentation could contain noise. To improve that, we have considered some specificities of our images and observed that the tip of mandible is less noisy than its base. So, first we have sorted the contour points according to their y-value to build the subset containing the half upper part of contour points. Second, we have enhanced the PCA by Iterating steps on this subset until the stabilization (PCAI). At each iteration, the rotation value is refined and the procedure will stop when the new computing angle is lower than 1.5 degrees (see Fig. 4).

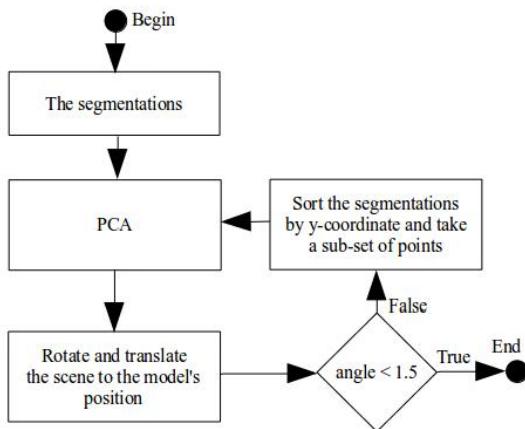


Figure 4: The workflow of the PCAI, allowing to refine the translation and the rotation values.

The Fig. 5 shows an example of the successive results obtained with PCAI. The red contours belongs to the model, the black one is the scene contours after one iteration, and the blue one is the resulting contours after iteration steps.

2.3 Landmark detection using SIFT

The last task of the workflow consists in estimating the landmarks of the scene from the manual ones of the model. SIFT method has been designed to identify points of interest inside images [10]. We have also used it to fix landmarks but with some modifications of the usual method. In order to reduce the computing time and the possible errors of location, we do not consider all points of the image and only take into account the area around each landmark on the model. Firstly, the region (called patch) around each landmark of the model is computed then extracted in the scene image at

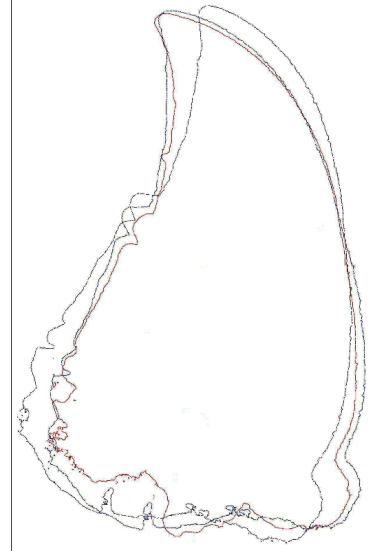


Figure 5: Iterations of the registration step between the model contour (in red) and the contours of the scene image.

the same position. Then, the SIFT descriptor is computed: the orientation and the gradient magnitude are calculated for each pixel by using the gradient values computed during the Canny step by applying the following equations (1):

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

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 gray value at position (x,y) in the image,
- $m(x,y)$ is the gradient magnitude of the pixel at position (x,y) ,
- $\theta(x,y)$ is the orientation of the pixel at position (x,y) .

The SIFT descriptor for each patch is an histogram containing the sum of pixel gradients for each considered direction. As usually, eight directions are taking into account ($0^\circ - 45^\circ$, $46^\circ - 90^\circ$, $91^\circ - 135^\circ$, $136^\circ - 180^\circ$, $181^\circ - 225^\circ$, $226^\circ - 270^\circ$, $271^\circ - 315^\circ$, $316^\circ - 360^\circ$). Finally, the feature vector is normalized to reduce the effects of illumination changes.

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

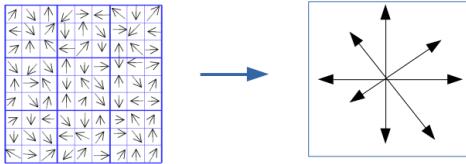


Figure 6: Calculus of the SIFT descriptor for a patch. In the right figure, the arrow length corresponds to the gradient value.

The comparison between two SIFT descriptors is done by using the L_2 -distance with the following equation (2):

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

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 values.

The Fig. 7 illustrates how we have applied SIFT into our workflow. To detect the scene landmarks, the patches P_m of the model and P_s of the scene are created with the size of P_m smaller than the size of P_s . After experiments, we have kept 36X36 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 the 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 position of the estimated 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 set at the original location on the original scene image by applying the reverse operation of rotation and translation.

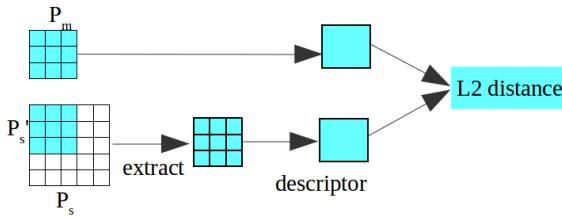


Figure 7: Steps of descriptors comparisons between the patch P_m of the model image and the patches P'_s of the scene image.

3 EXPERIMENTS AND RESULT

The complete method is implemented in the framework MAELab¹. The left and the right mandibles of the beetles has been analyzed separately. After verifying the

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

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 between 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.

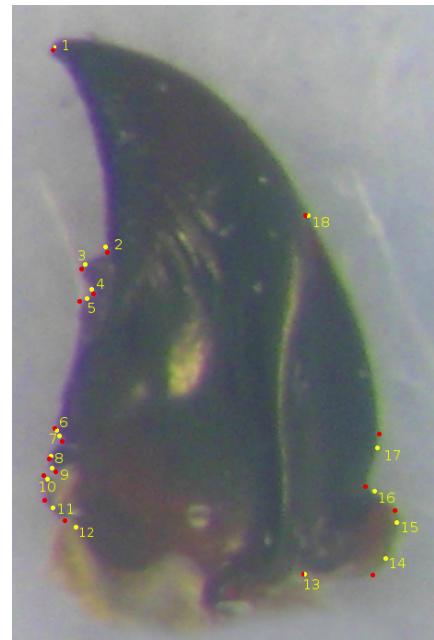


Figure 8: The manual (in red) and estimated (in yellow) landmarks on a right mandible.

The 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 be-

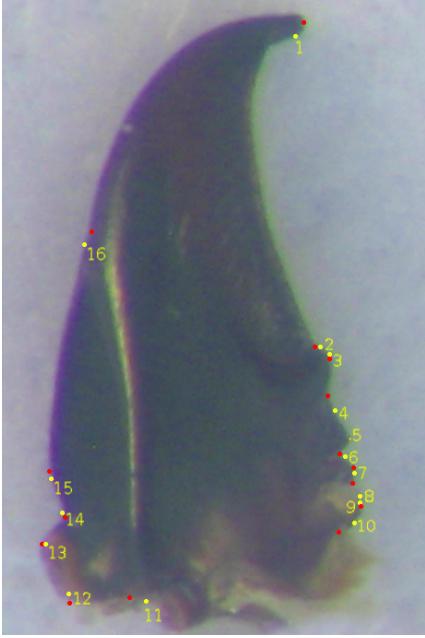


Figure 9: The manual (in red) and estimated (in yellow) landmarks of a left mandible.

tween 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.

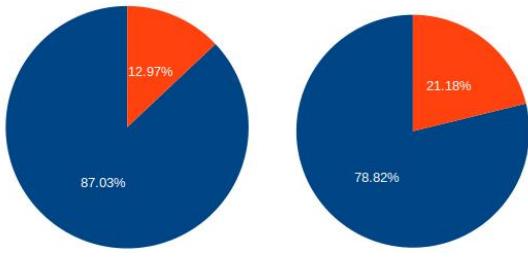


Figure 10: The mean proportion of well and bad landmark locations of the two sets of left and right mandibles.

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 1st estimated landmarks is very accurate with **98.62%**. The lowest proportion is **74.48%** for the 14th 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 1st landmark and **60.14%** for the 16th 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 11th and 12th landmarks of the left mandible and 13th and 14th 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.

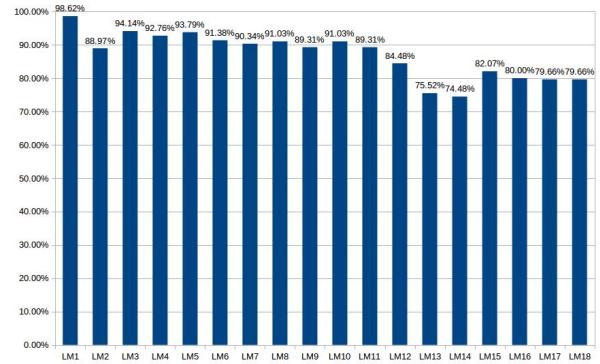


Figure 11: The proportion of well estimated landmarks of right mandibles.

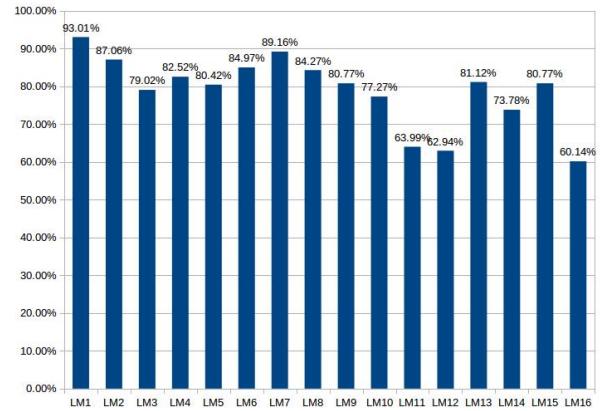


Figure 12: The proportion of well estimated landmarks of left mandibles.

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 an 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 use the SIFT descriptor to find the best matching position. The results show that the method succeed in locating the landmarks for all images. The accuracy of the method is sufficient to be proposed to biologists as 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 in 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 have rised up in image processing recently. Biologists are interested in 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 María 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] Dehua Li Jun Zeng. An adaptive canny edge detector using histogram concavity analysis. *International Journal of Digital Content Technology and its Applications*, 5, 2011.
- [8] L Lê Vănh, 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] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6):559–572, 1901.
- [12] Jonathon Shlens. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.
- [13] 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.