MAELab: a framework to automatize landmarks setting

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

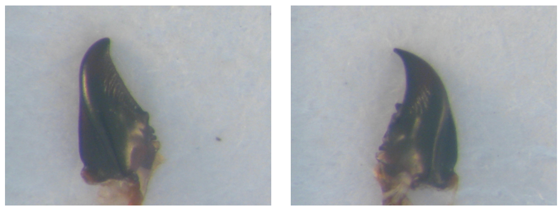
Phenotype of species are characterized by several informations like, age, sex, morphological measures and environment parameters. Biologists are familar to manually get morphological measures and in the case of analysis at the macro level (tissues, small parts of animal ...) they can do directly by measuring the element geometry: length, width, diameter, angles, ... Another way to obtain morphological measures is to take pictures of these elements and to run image processing algorithms. In order to evaluate a population of beetles into Brittany lands a collection of 293 beetles have been built. For each beetle, the images of the left and right mandibles are available and a set of 16 and 18 landmarks (resp. left and right) have been manually set, given to us the ground truth to compare to the estimated landmarks. In a previous work[4] based on Palaniswamy article [6] we have shown that if we consider the centroid measure of the mandible as the parameter to obtain, the probabilistic Hough Transform (PHT) can provide very interesting results. But if the goal is to consider more precisely the position or the geometry of landmarks area, the results are not accurate enough to consider estimated landmarks instead of manual landmarks. In this next round, we have modified and improved the previous algorithms of segmentation based on Canny algorithm, and we have left PHT and prefered a registration procedure based on an iteration of the Principal Component Analysis calculus. To achieve the estimated landmarks setting we have computed a SIFT descriptor limited to the landmarks area in the model and correlated it to the scene image. A workflow, MAELab, has been written, containing all the operations; it is available as library functions on a github website.

Keywords

Automatic morphology, landmarks identification, image registration.

1 INTRODUCTION

In biology, morphology analysis is widely used to keep the changing information of the organism or detecting the difference information between the organisms. From the result of morphology analysis, we can conclude the evolution of an organism family, or we may classify the organisms. Especially in agriculture, morphology is one of best ways to learn about the variations of the insect on crops. The morphology methods may be divided into the groups by the features which are used by the methods such as shape, structure, color, pattern or size of the object. In the aim to study the potential links between these variations and agricultual ecosystems, a set of 291 beetles has been collected with all the information about the sex, place where they are found and agricultural practices in each field were recorded. For each beetle, the morphometric landmarks has been defined on each part (each insect includes five parts: head, pronotum, body, left and right mandible) of the insect by the biologies. In this context, we try to indicate the landmarks on two parts of beetle (left and right mandible (see figure 1)). Morphometric landmarks are points that can be defined in all speciemns and located precisely. Landmarks are widely used in many biological studies and they are currently included into the classification procedures.



(a) Left mandible (b) Right mandible

Figure 1: The mandibles of beetle

In this paper, we focus on a method that can automatic identification of landmarks on 2D images of beetle, specify the mandibles of beetle. Through whole the article, we use two images to automatic extraction the landmarks: model image and scene image. The method mainly includes three stages: firstly, we extract the features of the object in the image; secondly, principal component analysis iteration is used to register two images, the translation and rotation between two images are also determined; finally, a refinement of the estimated landmarks is done by SIFT method.

In section 2, the steps of our methods will be presented. All experiments and evaluation are described in section 3.

2 METHOD

The problem to solve is to suppress the manual operation of setting landmarks on each image. To do that we propose a chain of operations, a workflow, including segmentation step of each image and a registration with a model image. The model image is more or less chosen randomly because we first verify by hand that it is not an image of some broken mandibles. The figure 2 shows the different steps of the workflow.

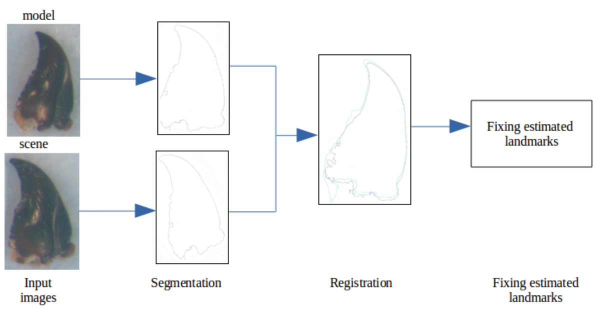


Figure 2: Overview of the proposed method

In this section, we will describe the different algorithms that we have used. It is worth to note that a protocol to take pictures of each mandible has been defined. All have been taken in the same conditions and the same camera with the same resolution.

2.1 Image segmentation

Segmentation is often the first task and the bootleneck of image processing chain. The most well-known algorithms are mainly classified as contours or regions based segmentation. We have chosen a contours one, the Canny alogrithm[2] allows to determine the list of edges belonging to the shape of the image. To use this method, two threshold values have to be set. As it is often mentioned, fixing the right values for these thresholds could be difficult[3]. The mandatory *threshold value* used by Canny algorithm has been determined by analyzing the image histogram (see [4] for detail). Most often authors define from this threshold, a lower and an upper one. The usual ratio of these two thresholds is $T_{lower} = (1/2) * T_{upper}$. In order to consider a larger range of values, we have prefer to set T_{lower} to $1/3$ of T_{upper}. For optimization purpose of the computing time, the gradient direction of each pixel which belongs to the contours is computed during the Canny algorithm step in order to be used later.

(a) Segmentation result after applying Canny algorithm (b) Segmentation result after applying Canny algorithm and post-process

Figure 3: The segmentation results of the image

To achieve the segmentation, the obtained contours are post-processed to remove unnecessary ones. As it is shown in figure 3, the final result of Canny generates some contours which do not belong to the shape of the mandible. With a simple algorithm, we browse the image and suppress the edges inside the main shape.

2.2 Image registration

As we have mentioned, all images have been captured following the same protocol at the same scale. But it remains differences in size between the mandibles (cause of size of the beetles), or orientation and position because of the mandible position under the camera. The next step concerns registration of model and scene before estimating the landmarks. Principal Component Analysis (PCA) is a well-known method to find the rotation and translation parameter values between two images. In our workflow, we have first used a classical PCA computing[7], [8].
As input values, we use the lists of points which has been defined by the segmentation step. Firstly, the centroid point and principal axis of each image are defined: the centroid point is the point which has the coordinate equal to the mean coordinate of all boundary points; the principal axis is a connected line from the centroid point to a point in the list of contours points which is determined as followed(algorithm 1):

The translation is indicated by the coordinate difference between the centroid points of the scene and the model.

Algorithm 1: The algorithm for finding the principal axis of a list of contour points

```
Input  : Centroid point, list points of contours
Output: The principal axis
for all points i in the list of contour points do
    Draw line l between i and centroid;
    for all points j in the list of contour points do
        if i ≠ j then
            Compute the perpendicular distance P_d
            between line l and point j;
        end
    end
    Compute the average (P_m) of all P_d;
    if P_m is minimal then
        Store i as i_min;
    end
end
The principal axis is the line between centroid and
i_min.
```

The rotation angle is the angle between the principal axes of these two images. The rotation direction is determined by checking coordinates of the principal axes (endpoints) the model and the scene. Then, the scene is moved to be register with the model. However, in some case, the translation and rotation between two images are not enough right because the result of the segmentation could be not perfect. To improve registration, we have enhanced the PCA by iteration steps (PCAI). We have considered some specificity of our images and observed that the tip part of the mandible is less noisy than the base. So, we sort the points according to their y-value. We build a subset of points which contains half part of points, these ones belonging to the upper part of the image. PCA is again completed for this subset to refine the rotation and translation values. This operation is iterated until the new computing angle is less than 1.5 degree (see figure 4).

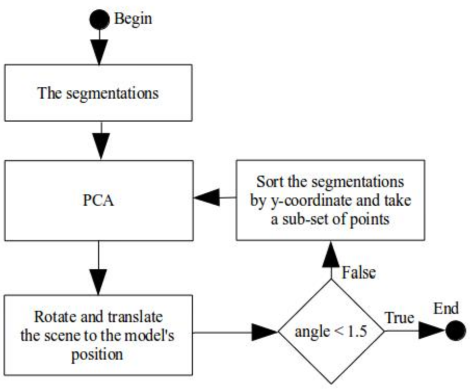


Figure 4: The flows in PCAI

Figure 5 shows an example of obtained results from the different steps of PCAI. In this figure, the red contours is the model segmentation, the black contours is the scene segmentation after one iteration, and the blue contours is the last result after finished PCAI .

Figure 5: Differnent registration steps between two images

2.3 Fixing the estimated landmarks

The last task in the workflows concerns how to estimate landmarks of the scene from the manual ones of the model. We have done by using SIFT[5] method. But we do not consider all points of the image as usually but only the area around the landmarks. Firstly, the region around each manual landmark (called patch) in the model is created and its corresponding position in the scene image is defined. Then, the SIFT descriptor is computed: the orientation and gradient magnitude are calculated for each their pixel by using the gradient values computed at the Canny step and applying the equation (1):

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

Where:

- $P(x,y)$ is the gray value at position (x,y) in the patch,
- $m(x,y)$ 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 which contains the sum of pixels gradient (computed at the Canny step) for each considering direction. As usually, eight directions are take into account ($0^o - 45^o$, $46^o - 90^o$, $91^o - 135^o$, $136^o - 180^o$, $181^o - 225^o$,

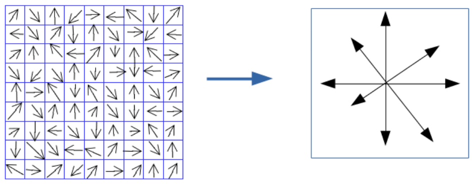


Figure 6: Calculus of the global descriptor for a patch. Gradient value is arrow length in the right figure

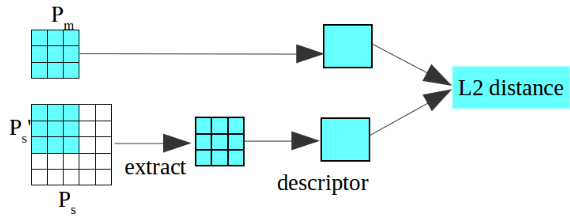


Figure 7: Illustrate the steps of descriptors comparison.

$226^o - 270^o$, $271^o - 315^o$, $316^o - 360^o$). Finally, the feature vector is normalize to reduce the effects of illumination changes.

The figure 6 shows a patch of 9x9 pixels created around each landmark on the model. The landmark of the model is located in the center of the patch. The size of 9x9 has been retained after several tests. Patch sizes: 18x18, 36x36, 54x54 have been also computed and gave us results statistically worst. From the histogram of the patch, we obtained the global gradient value for each direction.

The comparison between two SIFT descriptors is done by using $L2$ distance, 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 with size n,
- $D1_i$, $D2_i$ are the value at the location i in each descriptor.

To fix model landmarks on the scene, the patches P_m, P_s are created with size(P_m) $<$ size(P_s). After experiments, we keep 36 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 $P_s^{'}$ is not possible to get from P_s (border limits), the pixels outside $P_s$more considered. Then, the distance $L(P_m, P_s^{'})$ is computed by 2. This process is finished when all the pixels on the patch P_s are considered. The coordinates of estimated landmark is the location in P_s that has the smallest measure distance value with P_m. Finally, the coordinates of the estimated landmarks is set to the original location of the scene image by applying the reverse operation of rotation and translation.

3 EXPERIMENTS AND RESULT

All the steps in our method are implemented in MAELab[1]. The sets of beetle have been analyzed, right and left mandibles. After verifying the quality of the image, it remains 290 usable images for right mandible and 286 images for the left mandible. The removed images include the images that do not contain the mandible or the broken mandibles. In all valid images, a set of 18 manual landmarks of right mandible (16 landmarks for left mandible) are indicated by biologists. We have run the full workflow on all images. The preliminary results shown differences in algorithm performances, estimated landmarks on some scenes are well positioned and some another ones the precision is not satisfying. As we mentioned before, mandibles can exhibit different sizes because beetles have also. It seems that our method is sensible to this parameter. To improve our results, we have insert a step before computing SIFT descriptor to estimate the scale between one scene image and the model. The bounding boxes of the mandible of the model image and this one of the scene image are computed and the scales of x and y-direction are determined by the ratio between the corresponding sides of the bounding boxes. Then, the scene contours is scaled to fit the model curve.

Figure 8 and 9 show a complete result on one right mandible and one left mandible with the manual landmarks (red points) and estimated landmarks (yellow points). As we can see in the image, almost the estimated landmarks are quite near with the manual landmarks. This shows that our method is worked well for indicating the landmarks on the mandibles.

Figure **xxxxx** shows the global results that we have obtained **"xx%"** of estimated landmarks can be considered as well positioned because the distance between them and the manual landmarks are ground-truth.

Besides using the global experiments. We are also interested in the accuracy on the position of the estimated landmarks. In this experiment way, we calculate the distance between each manual landmark and corresponding automatic landmark. Through, we want to examine replacing the manual landmarks by corresponding estimated landmarks.

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

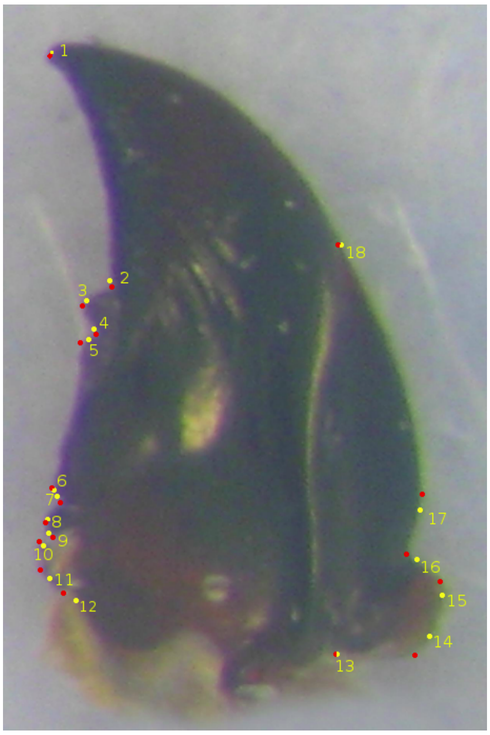


Figure 8: The manual and estimated landmarks on right mandible

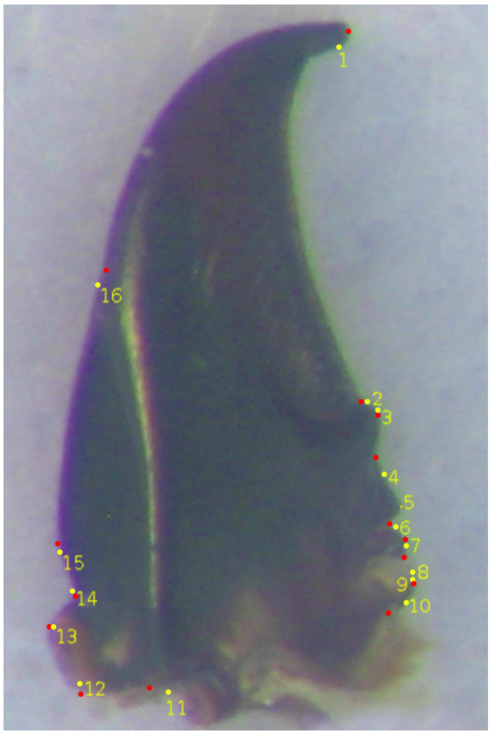
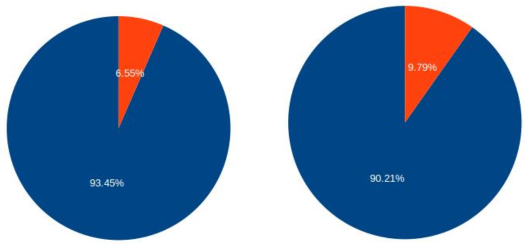


Figure 9: The manual and estimated landmarks on left mandible

Figure 11 and 12 show the correct proportion on each landmark of the mandibles. With 18 landmarks of right mandible, the position of the first estimated landmarks is reasonably accurate with 98.62%, the lowest proportion is 74,48% for fourteenth landmark. The remaining landmarks are also indicated with a high proportion (with the accuracy proportion greater than 75%). For left mandible, the highest and lowest success rate are 93,01% for the first landmark and 60,14% for the sixteenth landmark. The statistic is done on each automatic landmarks of all the images with a standard deviation[1]. As we can see in figure 3, the noisy of the base in segmentation is higher than the tip. This explains why the correct proportion on 11^{th}, 12^{th} landmark of the left mandible or 13^{th}, 14^{th} landmark of the left mandible are less than other landmarks. Moreover, when we reconsider the datasets, the images in left mandible are having more type of size than on the right mandible (scale problem). This is shows the reason that the success rate on right mandible is always greater than left mandible in both of the experiments.

From two experiment ways, we can see that the method is success in indicating all landmarks for each image; and the location of the landmarks is considered near with the manual landmarks in some aspect. In a different side, when comparing with our previous study (see in [4]). This method has more exactly about the position of automatic landmarks as well as the advantages



(a) Right mandibles

(b) Left mandibles

Figure 10: The correct proportions on centroid size of the mandibles

for the implementation process. The memory to detecting the landmarks, along with the times to execute the process are decreased dramatically.

4 CONCLUSION

Morphometric analysis is a powerful tool in biology in classification the species. Automatic identification

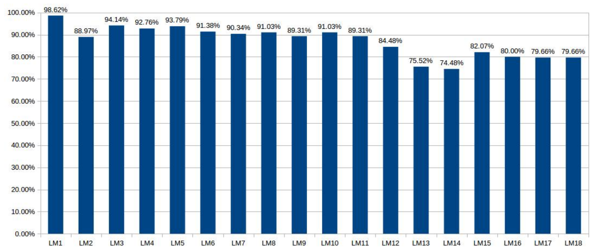


Figure 11: The correct proportions on each landmark of right mandibles

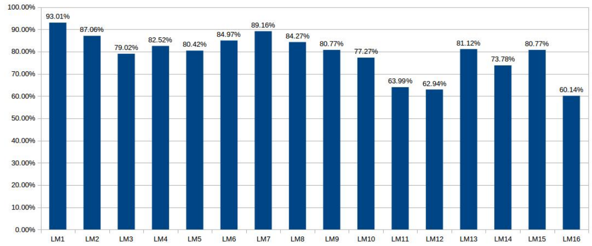


Figure 12: The correct proportions on each landmark of left mandibles

the characteristics biology of the organism is a difficult problem. In the content of this paper, we have begun to design a method to segment the beetle mandibles and to indicate automatically landmarks which have been determined by biologists. Each mandible is segmented by applying the Canny algorithm. Using PCAI to align the images. Finally, the estimated landmarks are indicated by applied SIFT descriptor. The first version of this method has been implemented. From now, the next stage of our method is to add the features to have the position of landmarks more precisely, i.e diagnose on the scale of the image.

5 REFERENCES

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

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

[3] Dehua Li Jun Zeng. An adaptive canny edge detector using histogram concavity analysis. *International Journal of Digital Content Technology and its Applications*, 5, 2011.

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

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

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

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

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