Automatic detection of individual and touching moths from trap images by combining contour-based and region-based segmentation

Mohamed Chafik Bakkay, Sylvie Chambon, Hatem A. Rashwan, Christian Lubat, Sébastien Barsotti

Automatic detection of individual and touching moths from trap images by combining contour-based and region-based segmentation

Mohamed Chafik Bakkay[1], Sylvie Chambon[1] ✉, Hatem A. Rashwan[1], Christian Lubat[2], Sébastien Barsotti[2]
[1]*INP-IRIT, University of Toulouse, IRIT, Toulouse, France*
[2]*SiConsult, Labège, France*
✉ *E-mail: schambon@enseeiht.fr*

Abstract: Insect detection is one of the most challenging problems of biometric image processing. This study focuses on developing a method to detect both individual insects and touching insects from trap images in extreme conditions. This method is able to combine recent approaches on contour-based and region-based segmentation. More precisely, the two contributions are: an adaptive *k*-means clustering approach by using the contour's convex hull and a new region merging algorithm. Quantitative evaluations show that the proposed method can detect insects with higher accuracy than that of the most used approaches.

1 Introduction

Many computer vision tools have been introduced in the literature in order to detect or to recognise objects in images, from the most ancient to the most recent: by using classic contour detection [1, 2] or snake contour [3], by clustering with *k*-means [4] or mean shift [5], and by exploiting key point detection like the scale invariant feature transform [6], to name a few. All these techniques have been applied to various industrial, medical or biometric applications [7], and also in the context of insect monitoring or detection [8], which is very crucial in agriculture. In fact, surveying insect species and evaluating their density in the fields allow farmers to forecast invasions of insects, and, consequently, to adapt the use of insecticides. Indeed, they can know when exactly insecticides, which are expensive and dangerous for plants and humans, can be used and, in consequence, it is possible to reduce the amount of product used.

Manual counting of insects from trap images is slow, expensive and sometimes error prone. Thus, developing a system which can achieve a completely automated detection, which can recognise and count insects is very advantageous. However it is a challenge. Actually, trap images may contain many types of noises: very small insects or herbs, the pheromone cap or some lines of glue (see Fig. 1 as an example). In addition, since the trap is installed in outdoor environment, the images are faced to illumination changes and techniques developed in a controlled environment are not adapted [9]. Finally, touching and overlapping insects can be found in the trap which will also complicate the counting task.

In conclusion, the problem of detecting and separating insects can be defined as the problem of segmenting a small object based on colour and shape characteristics in an non-homogeneous background that contains some difficulties. In this paper, the application will focus on a particular moth, however, for the proposed segmentation method, the size alone is taken into account, and it allows this segmentation to be as generic as possible for detecting any kind of moths or insects that have similar shape and size.

Some computer vision techniques are difficult to adapt to this task, or they are not enough efficient to be used alone. For example, image segmentation using active contours needs an initialisation that is close to the object and is not really easy to adapt to multiple objects detection. Using key point detection alone is also not adapted, because these points do not contain enough information to recognise and to separate insects (to capture the entire shape of the insect, it seems natural to have a lot of points on all the contour of the shape). In consequence, computer vision techniques used for insect monitoring based on images are: image restoration or enhancement (to take into account the presence of noises or artefacts), detection or segmentation (to separate the different elements), recognition based on unsupervised/supervised learning (to identify the nature of the elements).

In this paper, our purpose is to study the invasion of a particular moth, which is Lobesia Botrana (*Eudemis*), a European vine moth, for adapting the pesticide treatment of grape culture. More precisely, wine producers usually capture this particular moth, they count the number of insects and then they analyse the evolution of this counting in order to confirm the use of the pesticides or not. Consequently, the goal of this work is to introduce an automatic counting system of these insects in order to avoid the mobility of the wine producer and also to significantly reduce the use of pesticides. For that, we have to segment and to recognise insects captured inside a trap. However, in the proposed work, we will only focus on the segmentation step with two main contributions:

i. The proposition of *an adaptive k-means clustering* that is able to eliminate different types of noises, i.e. artefacts or non-insect elements, in the trap images. This approach is a preliminary accurate insect detection with their details (legs, antennas) using a recent robust contour estimation published in [10]. This method reduces the effects of illumination changes and light reflections. The adaptive part of our proposed algorithm is based on using the convex hull of these detected contours.
ii. The introduction of *a region merging algorithm* for separating touching insects.

The remaining part of this paper is structured as follows. Section 2 is devoted to a brief synthesis of the most relevant works on insect segmentation methods. In Section 3, the proposed method for insect segmentation is presented. In order to demonstrate the effectiveness of the proposed work, some experimental results and a comparative study are shown in Section 4. Section 5 concludes the paper and presents some directions for future works.

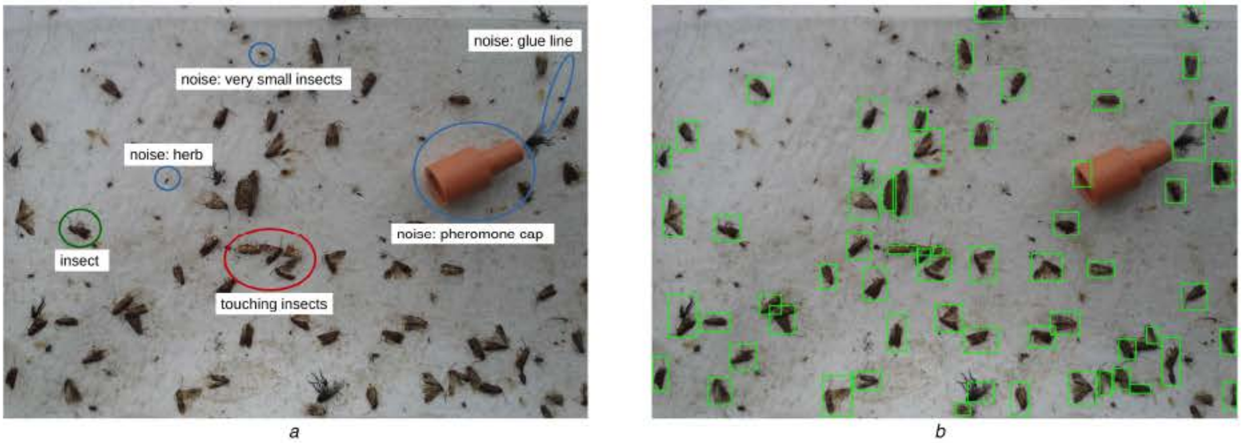


Fig. 1 *Illustration of the input images and the elements to detect*
(a) Input trap image: it contains the insects of interest, i e Lobesia Botrana, *Eudemis*, a European wine moth (inside the green circle) Unfortunately, it also contains some difficulties that we generally called noises, i e pheromone cap, herbs and small insects (blue circles) Finally, some of the insects of interest are too close to be separated (red circle), i e they are touching, and this problem has to be taken into account, ***(b)*** Results: it shows the kind of detection result that we have obtained All the insects have been individually detected More comments will be given in the experimental part, Section 4

2 Related work

In this section, most of the references are given in the field of identification and counting of insects but each time it is necessary, more general references of computer vision are given. As explained in the introduction, there are two possibilities for distinguishing and recognising elements on an image: to use one-step process based on a learning process for both detecting and identifying the elements or two-step process: to use a learning process or to segment the image and then to classify the elements. These two possibilities are developed in this section.

2.1 Learning-based approaches

In this category, existing methods have considered images captured with good poses, high resolution, clean background and under uniform lighting situations [11]. In this case, different classifiers are directly applied such as support vector machines [12–14], artificial neural networks [13, 15, 16], *k*-nearest neighbours [17, 18] and, recently, convolutional neural networks [19]. We have to remark that these methods first detect insects with a sliding window approach that is time consuming. In this paper, the acquisition conditions are not always as ideal as described in these existing approaches and this is why, in the rest of the paper and in the contributions, we will focus on proposing a first segmentation step, as robust as possible for localising insects in order to find initial detection candidates for a more accurate classification step.

2.2 Contour-based and region-based approaches

The literature is abundant in the field of segmentation and it is common to distinguish contour-based approaches [3] from region-based approaches [20]. However, some algorithms have also been inspired by classification-based methods, like mean shift [5]. In addition, more recently, over-segmentation techniques with superpixels have been introduced [21]. For insect detection and counting, only a subset of these approaches has been introduced in this domain: thresholding-based segmentation, clustering-based segmentation, contour-based segmentation and region-based segmentation.

Regarding *thresholding-based segmentation*, some approaches use histogram-based threshold [22, 23], adaptive threshold [24], fuzzy set entropy-based threshold [25] and the Otsu technique [26]. These methods are efficient when the images do not contain other different objects. For *clustering-based segmentation methods*, some approaches are using *k*-means technique [27], the fuzzy C-means clustering [28] or expectation-maximisation clustering [8]. These segmentation methods are generally fast and accurate, since they keep details of insects [29]. However, they strongly depend on parameters given as input, such as a number of classes and a threshold, which can change according to the type of image. Even automatic thresholding and clustering generally fail to separate touching insects, since a majority of methods depends on the colour information alone. In conclusion, clustering-based approach seems to be an interesting option for the objective of this paper, but only if it is adapted to better take into account not only the colour of the insects, but also the shape and the size of the insects of interest.

When using *contour-based segmentation* approaches, as the goal is to estimate the contour of the shape, we can expect to obtain more details of the shape of the insects compared to other segmentation techniques. However, most contour algorithms such as Canny [1] suffer from discontinuous contours and are still affected by noises. To alleviate these difficulties, some methods used active contours, such as snake [30]. However, an initial contour is necessary and it makes this approach difficult to use in an automated context. Contour-based methods also fail to separate touching insects, since only one contour for all touching insects is detected. Again, we can conclude that contour-based methods are interesting for this task, but only if they are improved to be more robust to noises and if they are combined with other approaches, in particular to separate touching insects.

Finally, most of *region-based segmentation* approaches are able to separate touching insects, like watershed method [31] and graph-cuts optimisation [32]. However, these techniques suffer from over-segmentation. In both cases, it is mostly due to the difficulty to select the seeds needed by the propagation or optimisation step of the algorithm and in most approaches, it is chosen to take local maxima, so to have more seeds than the real number of regions to segment. Hence, the same conclusion is made: it is needed to combine this kind of segmentation approaches with a complementary method, like contour-based one, to cope with the over-segmentation problem.

In consequence, in the literature, *hybrid segmentation methods* have been introduced. They combine multiple segmentation algorithms. For instance, Mele [33] have proposed an algorithm in two steps: a coarse-global segmentation and a fine-local segmentation. The coarse-global step contains a global thresholding method that eliminates small objects. In the fine local, they have applied a seed region growing algorithm. The crucial part of this algorithm is that it requires an input of seeds for background and foreground, which can be difficult to automate. Wen *et al.* [11] have introduced a two-level morphological-based method. For the first level, like [33], a global thresholding method is applied to eliminate small objects. For the second level, they use morphological operators, which need a parametrisation depending on the type of the input image. Finally, the approach of Yalcin [34] consists in first applying a background subtraction based on Gaussian mixture models, and then using an active contour model in order to extract the pixels that belong to the boundary. However,

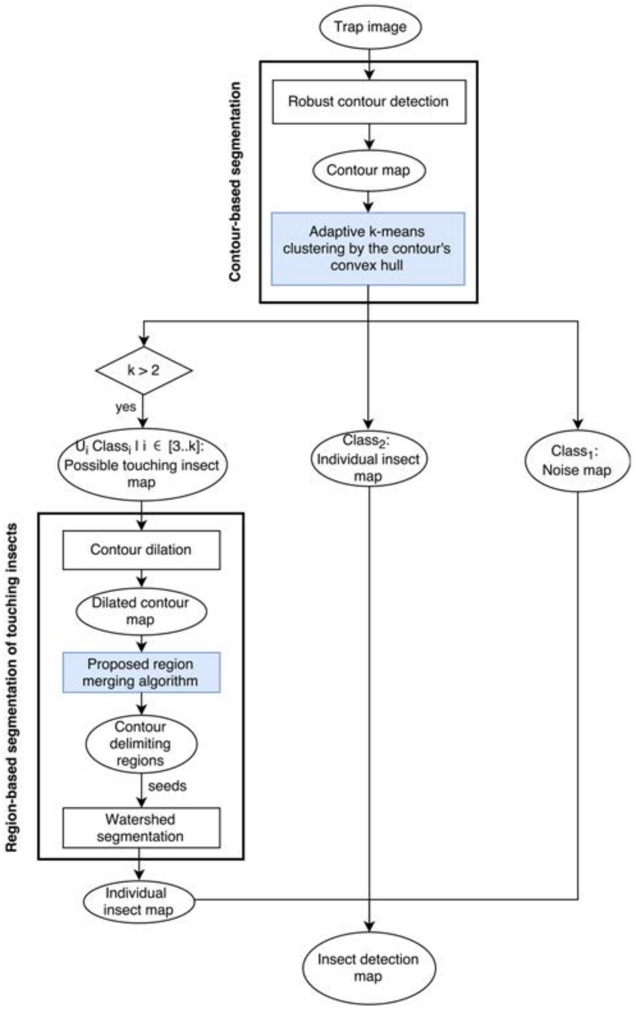


Fig. 2 *Proposed method – we distinguish the manipulated data (inside ellipses) from the actions (inside rectangles). The main contributions of the paper correspond to blue rectangles*

this method needs multiple images taken at different times per day and a background model in order to separate touching insects.

2.3 Discussion

In conclusion of this review, it seems that most of the existing approaches are incomplete, since they take into account only one aspect of the problem (the colour or the shape, the contour or the region). In consequence, they fail to take into account some difficulties like the presence of touching insects. In addition, even the most recent hybrid approaches introduce many parameters not easy to choose. Consequently, our idea is to combine contour-based and region-based segmentation approaches to keep the details of insects, in addition to be able to separate touching insects. Thus, an automated segmentation method is proposed that can be enough reliable to initialise candidates for insect identification. Moreover, we want to introduce an approach with less parameters as possible and easy to choose, i.e. the choice of these parameters do not dramatically influence the quality of the results.

3 Proposed approach

3.1 Overview

The proposed method allows the detection of individual and touching insects in images of a trap, that contain many difficulties (noises and elements that are not insects), as presented in Fig. 1, by combining contour-based and region-based segmentation approaches. The method takes as an input a trap image and it returns as an output the localisation of each insect (i.e. a bounding box of each detected insect).

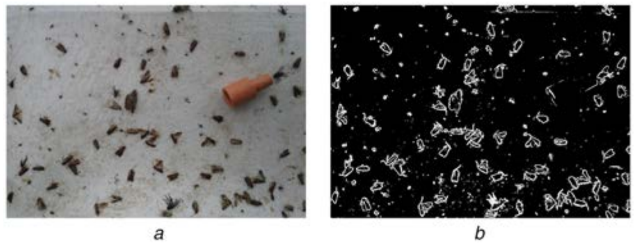


Fig. 3 *Contour detection*
(a) Colour image, ***(b)*** Estimated contours

Fig. 2 shows the schematic overview of the proposed method. First of all, we apply a *robust contour detection* that we previously published in [10] to detect the different contours in the input image. Then, we apply a *k*-means algorithm to classify the previous estimated contours into different categories. On this step, our main contribution is to introduce an adapted criterion: the shape of the surface included in a closed contour. In fact, the shape is a significant characteristic to separate the different kinds of elements in the scene, i.e. it helps to distinguish between contours due to noises (class 1), contours related to individual insects (class 2) and contours that contain touching insects (class 3 to *k*). Moreover, in comparison to the state-of-the-art methods, in this approach, the number of classes is automatically selected by using the Elbow method [35]. After this automated clustering step, the next task attempts to separate the obtained possible touching insects by applying a region-based segmentation. The idea is to use the contours classified into class 3 to *k* that delimit regions as seeds for the watershed algorithm [36]. This region-based segmentation part contains three ordered steps: the *contour dilation*, the *region merging algorithm* and the *watershed segmentation*. After these three steps, two results are possible:

i. The watershed algorithm detects two or more regions inside the contour, thus touching insects will be separated to two or more insects.
ii. The algorithm detects only one big region and the shape of this insect is just refined.

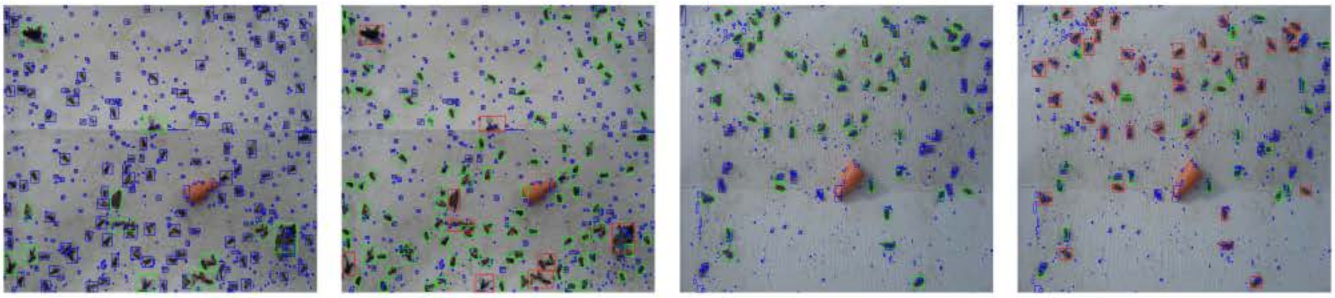
The details and the justifications of each step of the proposed scheme are given in the following subsections.

3.2 Robust contour detection (first step of the contour-based segmentation)

In the literature, many contour detectors have been introduced and the most famous one is Canny operator [1]. More recently, in [10] the interest of using curvature has been highlighted, since it detects what we have named ‘curvilinear structures’ that generate a single response for both lines and edges. Moreover, in [10], we have shown that using curvature allows to deal with noises. For these reasons, we use this detector in this work. In details, since the principal curvatures of a curve at a given point can be approximated by the eigenvalues of the Hessian matrix, the Hessian matrix can be used for estimating these principal curvatures. Then, we compute the difference between these principal curvatures (i.e. eigenvalues) and we suppose that the higher the difference, the most interesting the point, i.e. the point is related to an edge or a line. This computation is done in multi-scale in order to detect both important structures and small details. We have to choose the number of scales, N_S, and the choice for this parameter are given in Section 4.2. Finally, the curves are given by selecting the local maxima in scale space. An example of these estimated curves is shown in Fig. 3. More details of the approach can be found in the complete description in [10].

3.3 Adaptive k-means clustering (second step of the contour-based segmentation)

In this paper, we propose to use the contour's convex hull, noted $\mathscr{H}$, as a criterion for estimating the clustering. To illustrate the interest of this criterion, we introduce it in a *k*-means method



a b c d

Fig. 4 *Illustration of the behaviour of the Elbow method – for understanding the coding colours, see the beginning of Section 4. In (a) and (b), we have a first example where the Elbow is reached for k = 3, whereas in (c) and (d), it is a second example where Elbow is reached for k = 2*
(a) $k = 2$, ***(b)*** $k = 3$ (Elbow), ***(c)*** $k = 2$ (Elbow), ***(d)*** $k = 3$

a b c d

Fig. 5 *Region-based segmentation of touching insects*
(a) Colour subimage that contains multiple insects, ***(b)*** Contour detection, ***(c)*** Contour dilation and region merging, ***(d)*** Watershed algorithm We can see that the two insects have been correctly separated

because this approach is used intensively in the literature and it has proved its effectiveness. Moreover, it is the easiest method for introducing this criterion. However, this criterion can be incorporated in any other existing segmentation approach.

The main difficulty of using k-means approach is to choose the number of classes k. To try to alleviate this problem, we have studied the use of the mean-shift algorithm, which is a kind of k-means algorithm without choosing the number of classes. In fact, two parameters must be chosen: the scale in space and the scale in colour. These parameters depend indirectly on the choice of the number of classes. Frequently, in trap images, there are three main classes related to the object's size: noises (very small size, like small insects, herbs or lines of glue used to capture the insects), individual insects (i.e. having a medium size) and touching or big insects (i.e. having a big size). It seems natural to choose $k = 3$ but, actually, we can have a trap image without touching insects and in this case, there are only two classes (noise and individual insects). In addition, a trap image can contain many touching insects in different sizes, such as in Insect Soup Challenge dataset [33]. In this case, the group of touching insects contains elements that differ by the size but they cannot correspond to big insects. Consequently, the number of categories can be more or less than three. In conclusion, it is obvious that k cannot be always the same and have to be adapted to the input image.

To estimate the suitable number of classes, we used the 'Elbow' method [35]. It consists in starting with $k = 2$. Then k is increased by 1 and a distortion cost (DC) is estimated for each value of k, by using the formula

$$\mathrm{DC}(k) = \sum_{i=0}^{i=N_s} \| \mathscr{H}(c_i) - \mathscr{M}_{l_i} \|^2 , \qquad (1)$$

where N_s is the number of samples, i.e. the number of closed contours, noted c_i that have been detected, $\mathscr{H}(c_i)$ is the convex hull of each sample, i.e. each closed contours c_i, l_i is the array that stores the cluster index, i.e. the class whose each closed contour belongs to, and $\mathscr{M}$ is the array of cluster/class centres. When k increases, the cost goes down rapidly, and at a given value of k, the cost slowly decreases. It is said that, at this given value, the curve reaches an Elbow (see the experiment results, Section 4.1 and Fig. 4, to illustrate how the algorithm works). After estimating the best value for k, we suppose that the first class represents the noise, the second class represents the individual insects and all remaining classes represent possible touching insects. In the next step, the proposed method tries to separate these touching insects.

3.4 Region-based segmentation of touching insects

3.4.1 Introduction: Some methods, such as [37], use a segmentation based on optical flow and N-cuts algorithm. However, these techniques need at least two images of the trap to compute the optical flow. We do not have this kind of images in our application. The most used algorithm for separating touching objects is the watershed algorithm [36]. Some methods use a watershed algorithm based on mathematical morphology for detecting pests [31]. In [38], the author uses watershed segmentation based on prior information about the characteristics of the elements to segment. However, in the case of complex image, watershed still suffers from over-segmentation and strongly depends on the prior information.

Thus, our idea is to use the contour delimiting regions as seeds for the watershed algorithm to separate insects accurately while avoiding over-segmentation. More precisely, the proposed segmentation, presented in Fig. 5, contains these three ordered steps: contour dilation, region merging algorithm and watershed algorithm.

3.4.2 Contour dilation: To extract contour delimiting regions we should ensure that contours are closed. Although the step described in Section 3.2 provides accurate contours, the detected contours can contain some discontinuities. Thus, a morphological dilation of the contours is needed. In fact, each contour is iteratively dilated by two-by-two kernel until the area of regions inside this contour will be equal to S% of the contour's convex hull. We empirically select this threshold of dilation S in order to ensure that the contour will be as closed as possible. Section 4.2 gives explanations about the choice and the influence of this parameter. However, some isolated small regions inside this contour can appear. As an example, in Fig. 5c, we have the dilated contour of touching insects. The white big regions inside this contour represent the possible individual insects, while the white small regions represent parts of these

$\mathcal{B}$: set of points belonging to the background
N_m: Number of merges that have been made

procedure REGION −MERGING
  for each contour $c \in \mathcal{C}$ **do**
    do
      $N_m \leftarrow 0$
      if $(\exists r_i \in \mathcal{C} \mid \mathcal{A}(r_i) \leq T \times \mathcal{H}(c))$ **then**
        Find r_n: nearest region to r_i
        Compute l: shortest segment between r_n and r_i
        if $(l \cap \mathcal{B}) = \emptyset$ **then**
          Merge r_i and r_n : $r_i = r_i \cup l \cup r_n$
          $N_m \leftarrow N_{m+1}$
        end if
      end if
    while $N_m \neq 0$
  end for
end procedure

Fig. 6 *Algorithm 1 proposed region merging algorithm*

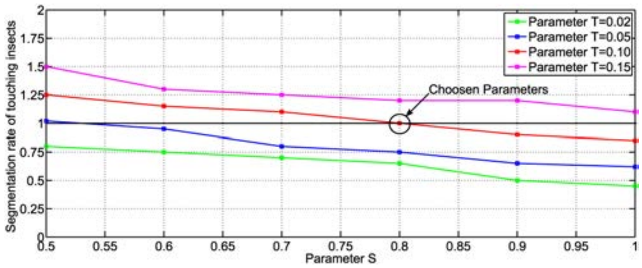


Fig. 7 *Parameter study using 20 images containing 58 groups of touching insects. We show how S and T have to be fixed in order to have a segmentation rate of touching insects equals to 1. All the definitions and explanations about S, T and this rate are given in Section 4.2*

individual insects. If we keep these small regions as seeds, the watershed algorithm will over-segment the insects in these regions. Thus, these regions must be merged with the bigger regions to avoid over-segmentation. This is the goal of the next step of the algorithm.

3.4.3 Proposed region merging algorithm: The proposed region merging algorithm, see Algorithm 1 (in Fig. 6), iteratively merges small regions r_i, inside each dilated contour c belonging to the set of estimated contours $\mathscr{C}$, with the nearest region r_n inside the same contour. Merging two regions consists of linking them by the shortest segment l between them. This segment has at least one pixel of thickness. An important fact is that the algorithm must avoid to merge big regions together (to keep insects separated) and must avoid to open the closed contours. Thus, the step of merging is repeated until no small regions, which can be merged, are found. In this way, small regions will grow progressively until they reach a maximum size, i.e. $\mathscr{A}(r_i) \leq T \times \mathscr{H}(c)$), where $\mathscr{A}(r_i)$ means the area of r_i and T is a threshold that is empirically selected, see Section 4.2 to illustrate the influence of this parameter. This condition avoids the algorithm to merge big regions. In Fig. 5*b*, we can see that the estimated contour contains some discontinuities, see for example the wings of the fly. After the dilation (see Fig. 5*c*), the contours are closed but some small regions still remain, like the region of the legs of the fly. However, after the merging step, as expected, we can notice that the small regions are merged with the biggest nearest regions, see, for example, the white path between the legs and the core of the fly, while the two insects are not merged together.

3.4.4 Watershed algorithm: Finally, after merging the small regions, the watershed algorithm [36] is applied with these merged regions as seeds in order to obtain an accurate segmentation, i.e. an accurate separation of insects. The result given in Fig. 5*d* highlights the quality of the results that we have obtained in general on all the tested images.

4 Experimentation

For all the results presenting in this section, we use these coding colours:

- Green rectangles are insects.
- Blue rectangles indicate noises (herbs, small insects, reflections).
- Red rectangles correspond to touching insects.

4.1 Illustration of the behaviour of Elbow algorithm

In Fig. 4, we present the results of the Elbow algorithm for two different images, and with two values of k. More precisely, for the first image, in Figs. 4*a* and *b*, the Elbow is reached for $k = 2$ whereas, for the second one, in Figs. 4*c* and *d*, it is reached for $k = 3$. These examples illustrate how it is important to use an adaptive number of classes and the interest of using the Elbow algorithm.

4.2 Parameter study

In the proposed algorithm, we have to choose these parameters:

- N_S*:* the number of scales used for the contour detection;
- *S:* the threshold used for the dilation step;
- *T:* the threshold used for the merging step.

For the contour detection algorithm [10], four scale levels are used for N_S, it is a choice coherent with the recommendations made in the paper. For the two other parameters, a value has been chosen empirically by analysing the graphs presented in Fig. 7. More precisely, to choose the values of S and T, we have used what we call the segmentation rate of touching insects which is equal to the number of segmented insects divided by the number of original insects (ground truth). When, this rate is above 1, it means that the image is over-segmented, whereas when it is under 1, the image is under-segmented. Hence, ideally, it has to be equal to 1, and, in this configuration, it means that the number of segmented insects is equals to the number of original insects. In consequence, in the graph, we have to choose the point of intersection of any curve with the black line (that corresponds to the segmentation rate of touching insects equals to 1) to avoid under- and over-segmentations while minimising the dilation (i.e. minimising the computational cost). For each curve presented in Fig. 7, T has a fixed value whereas S varies from 50 to 100% of the contour's convex hull. We show the curves for four different values of T between 0.02 and 0.15. Finally, the graphs highlight that the best choice, i.e. the choice that allows to obtain a segmentation rate of touching insects equals to 1, is $S = 80\%$ of the contour's convex hull and $T = 0.1$. Consequently, we define small regions as regions that have at least 10% of the contour's convex hull.

4.3 Real dataset

We have collected a big number of insect images (almost 100 images with an average of 30 insects per image; see Figs. 8*a* and *b*). These images are collected using moth traps designed by SiConsult, one of the companies involved in the project (see acknowledgment for the details). In these images, there are many insects of varying types and sizes and they are captured under different illumination conditions. These images contain different noises or elements that can induce false detections, like the pheromone cap (see Figs. 8*b*). As shown, most individual insects are detected by the proposed method. In addition, the proposed method separates touching insects in most of the cases. As well, it avoids over-segmentation of big insects (see Figs. 8*a*). However, in some images, such as the image shown in Figs. 8*b*, the big insects are over-segmented. Since some big insects have different parts (like big wings, or thin long paws), the watershed algorithm considers them as overlapping insects. However, we can imagine that the recognition step will not recognise these parts of insects as the insect we want to recognise because the shape is too different.

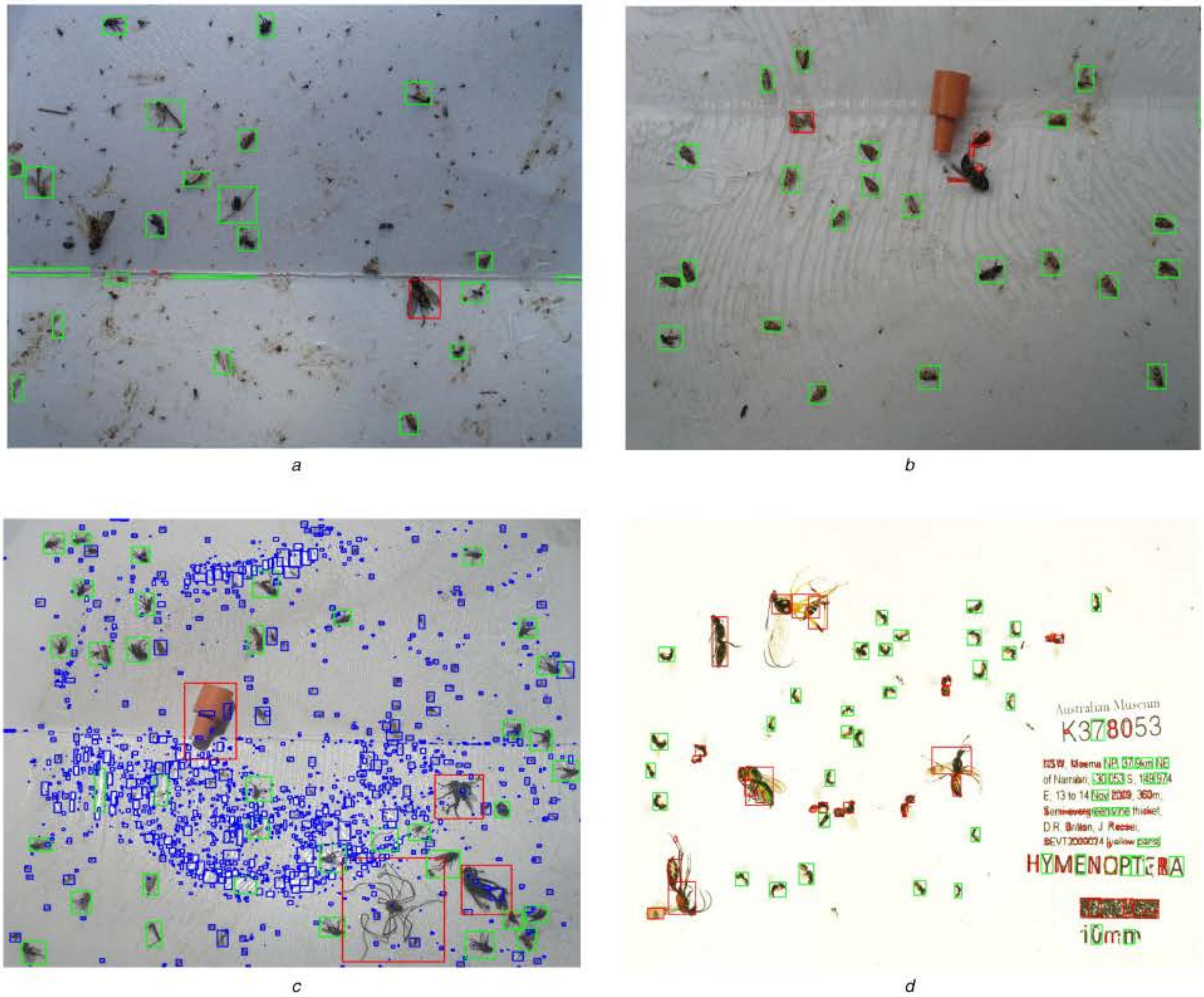


Fig. 8 *Results with real trap image dataset – for understanding the coding colours, see the beginning of Section 4. (a)-(d) are outputs of the proposed method using different input trap images. In (a) and (b), images are relatively simple on the contrary of (c) and (d). Most of the insects have been correctly detected and noises are removed, except in (b) that shows some errors due to the artefacts of the trap. It is important to notice that the proposed algorithm does not separate insects when it is not needed, see the fly on (a), which is an adapted behaviour for the recognition step. In (c) and (d) that are very challenging, the proposed method still detects and separates correctly the different insects*

Table 1 Confusion matrix

	Noise	Insects	Touching insects
noise	0.87	0.13	0
insects	0.13	0.70	0.17
touching insects	0	0.18	0.82

In Table 1, we present the confusion matrix [39] of the detection algorithm. It allows the visualisation of the performance of the proposed method by giving an idea about the distribution of the detection errors. In fact, all correct detections are located in the diagonal of the table: 87% of noises are detected as noises, 70% of insects are detected as insects and 82% of touching insects are detected as touching insects. The errors are represented by values outside the diagonal. We can see from the matrix that the proposed method classify 13% of noises as insects, however, no noise is detected as touching insects. In addition, 13% of insects are detected as noises, in turn 17% of insects are detected as touching insects. Finally, 18% of touching insects are detected as insects and no touching insects are detected as noise. Thus, the proposed method has small confusion for distinguishing between: (insects and touching insects) and (insects and noises). However, as expected, it can properly discriminate between noises and touching insects.

4.4 Details about the results with extreme conditions

We also tested the proposed method on images that contain noises and lighting defaults, i.e. light reflections. In these images, the noise is more important than the insects: herbs, very small insects or some parts of insects. In addition, the lines of glue are detected as contours. However, the proposed method avoids them and classifies them as noises. An example is given in Fig. 8*c*.

Finally, we tested the proposed method on images from the Insect Soup Challenge dataset used in [33], such as the image shown in Fig. 8*d*. With this dataset, we have selected images that contain insects that are close together and that can be touching or overlapping. Moreover, the images chosen contain noises (such as broken wings and other insect parts) (i.e. 9 images over 19 images). Only one constraint has to be respected and this is why we have kept only nine images: insect may have almost the same size. As shown in Fig. 8*d*, the proposed algorithm can correctly segment most of the individual insects, while it fails in separating some cases of touching insects. However, in general, we obtained an average detection rate of 82% with these images.

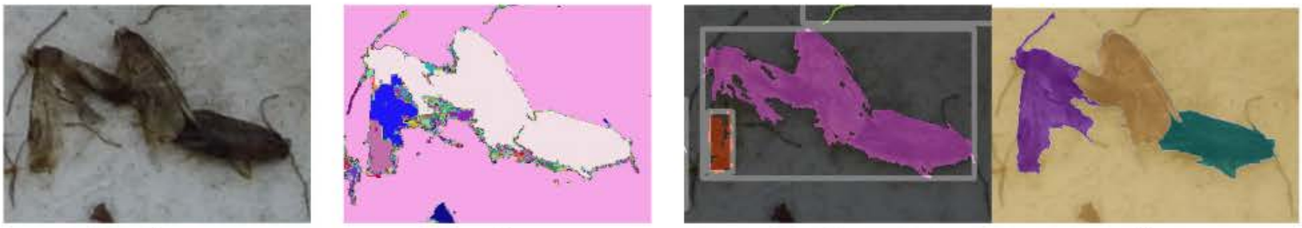
4.5 Comparison with existing approaches

Many approaches have been presented in Section 2, and we have chosen to compare the proposed approach with:

Table 2 Precision, under- and over-insect segmentation rates – comparison with six existing approaches

Method	Precision rate	Rate of under-segmented insects	Rate of over-segmented insects
Otsu [11]	0.23	0.69	0.08
Otsu [11] + watershed [31]	0.47	0.15	0.38
superpixel [21]	0.23	0.69	0.08
meanshift [5]	0.54	0.15	0.31
K-means [27]	0.31	0.61	0.08
graph cuts [32]	0.38	0.54	0.08
proposed method	**0.77**	**0.15**	**0.08**

Bold letters indicate the best results



a b c d

Fig. 9 *Comparison with the state-of-the-art*
(a) Colour image, ***(b)*** Mean shift [5], ***(c)*** Graph cuts [32], ***(d)*** Proposed method

- Some references in the domain of insect detection from the simplest one to the most sophisticated one: Otsu thresholding [31], Otsu thresholding combined with watershed [31] and an approach based on graph cuts [32];
- Some references that are the most famous in the domain of image classification and segmentation: mean shift [5] and superpixels [21].

It is important to notice that the most comparable approach to the proposed algorithm is the work of [31], since it is both dedicated to insect detection and based on the watershed algorithm.

In Table 2, for all the tested methods, the average detection rates for all tested images have been computed. We distinguish the precision rate, the rate of under-segmented insects and the rate of over-segmented insects. The proposed method obtains the higher precision rate with the minimum over- and under-insect segmentation rates compared with the existing methods. Mean-shift method also has a precision rate greater when 0.5 but significantly lower than the proposed approach. However, mean shift suffers from a high rate of under-segmented insects. Otsu, superpixels, *k*-means and graph-cuts methods yield low precision rate and it is related to the fact that they have a high rate of under-segmented insects.

In Fig. 9, we present visual results obtained for the three methods that yield the best average detection rate: the mean-shift method, the graph-cuts method and the proposed approach. The example contains three overlapping similar moths with similar shape but different colours. Regarding the visual results in details, the method based on graph-cuts contains less noises on the detected contours than the mean-shift algorithm. However, the proposed method seems to segment the details of the insects the most accurately. Second, the proposed method properly separates the three touching insects, while mean-shift and graph-cuts failed to correctly separate them.

5 Conclusion

In this paper, a new automated method has been proposed for detecting individual and touching insects from trap images. This method detects insects accurately with their details using a recent contour estimation that is robust to illumination changes. It is also able to eliminate different types of noises from trap images, since it is based on a proposition of an adaptive *k*-means clustering. Finally, it separates touching insects using a proposed region-based segmentation algorithm. Some quantitative evaluations showed that the proposed method can detect insects with higher accuracy than the accuracy obtained with six other existing approaches. For future works, the detection and localisation will be completed by adding a learning based recognition step. For that purpose, we will introduce a descriptor that can take into account colour, shape and size in order to be as discriminant as possible. Moreover, after validating the approach with the Lobesia Botrana butterfly, we plane to generalise it to other species.

6 Acknowledgment

This work was conducted as part of the project VITI OPTIMUM 2.0. It aims to develop several innovative diagnostic tools implementing ICT and IOT technologies (more precise, interoperable and connected) to facilitate and optimise the work of the technician of agricultural cooperatives on the vineyard parcel and allow to provide a personalised advice to the winegrowers. This is to reduce the use of phytosanitary products and to allow a technician to monitor 750–800 ha/year. The project brings together seven partners: a cooperative: QUALISOL; three companies: SICONSULT, SOGEFI, ISAGRI; two public laboratories: EI PURPAN, CLLE-LTC; a technical centre: IFV and two major subcontractors: Arboritech and IRIT. The project has been approved by the French competitiveness clusters: Agri Sud-Ouest Innovation and IAR. The project has been cofinanced by BPIFrance, European Regional Development Fund and the Occitania Region.

EUROPEAN UNION
PROJET IS PART-FINANCED BY
THE EUROPEAN REGIONAL DEVELOPMENT FUND

7 References

[1] Canny, J : ‘A computational approach to edge detection’, *IEEE Trans. Pattern Anal. Mach. Intell.*, 1986, **8**, (6), pp 679–698

[2] Deriche, R : ‘Using Canny's criteria to derive a recursively implemented optimal edge detector’, *Int. J. Comput. Vis.*, 1987, **1**, (2), pp 167–187

[3] Kass, M , Witkin, A , Terzopoulos, D : ‘Snakes: active contour models’, *Int. J. Comput. Vis.*, 1988, **1**, (4), pp 321–331

[4] Wagstaff, K , Cardie, C , Rogers, S , *et al.*: ‘Constrained k-means clustering with background knowledge’ Int Conf on Machine Learning, 2001, pp 577–584

[5] Comaniciu, D , Meer, P : ‘Mean shift: a robust approach toward feature space analysis’, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2002, **24**, (5), pp 603–619

[6] Lowe, D : ‘Distinctive image features from scale-invariant keypoints’, *Int. J. Comput. Vis.*, 2004, **60**, (2), pp 91–110

[7] Deschamps, T , Cohen, L : ‘Fast extraction of minimal paths in 3D images and applications to virtual endoscopy’, *Med. Image Anal.*, 2001, **5**, (4), pp 281–299

[8] Chen, X M , Geng, G H , Zhou, M Q , *et al.*: ‘Applying expectation-maximization in insect image segmentation using multi-features’, *Comput. Appl. Softw.*, 2009, **26**, (2), pp 20–22

[9] Schurischuster, S , Zambanini, S , Kampel, M , *et al.*: ‘Sensor study for monitoring varroa mites on honey bees (apis mellifera)’ Visual Observation and Analysis of Vertebrate and Insect Behavior Workshop, 2016

[10] Rashwan, H A , Chambon, S , Gurdjos, P , *et al.*: ‘Towards multi-scale feature detection repeatable over intensity and depth images’ IEEE Int Conf on Image Processing, 2016, pp 36–40

[11] Wen, C , Wu, D , Hu, H , *et al.*: ‘Pose estimation-dependent identification method for field moth images using deep learning architecture’, *Biosyst. Eng.*, 2015, **136**, pp 117–128

[12] Larios, N , Soran, B , Shapiro, L G , *et al.*: ‘Haar random forest features and SVM spatial matching kernel for stonefly species identification’ IEEE Int Conf on Pattern Recognition, 2010, pp 2624–2627

[13] Wang, J , Lin, C , Ji, L , *et al.*: ‘A new automatic identification system of insect images at the order level’, *Knowl.-Based Syst.*, 2012, **33**, pp 102–110

[14] Wen, C , Guyer, D E , Li, W : ‘Local feature-based identification and classification for orchard insects’, *Biosyst. Eng.*, 2009, **104**, (3), pp 299–307

[15] Fedor, P , Malenovský, I , Vanhara, J , *et al.*: ‘Thrips (thysanoptera) identification using artificial neural networks’, *Bull. Entomol. Res.*, 2008, **98**, (5), pp 437–447

[16] Kaya, Y , Kayci, L : ‘Application of artificial neural network for automatic detection of butterfly species using color and texture features’, *Vis. Comput.*, 2014, **30**, (1), pp 71–79

[17] Li, X L , Huang, S G , Zhou, M Q , *et al.*: ‘KNN-spectral regression LDA for insect recognition’ Int Conf on Information Science and Engineering, 2009, pp 1315–1318

[18] Wen, C , Guyer, D : ‘Image-based orchard insect automated identification and classification method’, *Comput. Electron. Agric.*, 2012, **89**, pp 110–115

[19] Ding, W , Taylor, G W : ‘Automatic moth detection from trap images for pest management’, *Comput. Electron. Agric.*, 2016, **123**, (C), pp 17–28

[20] Vincent, L , Soille, P : ‘Watersheds in digital spaces: an efficient algorithm based on immersion simulations’, *IEEE Trans. Pattern Anal. Mach. Intell.*, 1991, **13**, (6), pp 583–598

[21] Achanta, R , Shaji, A , Smith, K , *et al.*: ‘SLIC superpixels compared to state-of-the-art superpixel methods’, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012, **34**, (11), pp 2274–2282

[22] Blasco, J , Gómez-Sanchís, J , Gutierrez, A , *et al.*: ‘Automatic sex detection of individuals of ceratitis capitata by means of computer vision in a biofactory’, *Pest Manage. Sci.*, 2009, **65**, (1), pp 99–104

[23] Zhang, H T , Hu, Y D , Qiu, D Y : ‘The stored-grain pest image segmentation algorithm based on the relative entropy threshold’, *J. North China Inst. Water Conservancy Hydroelectr. Power*, 2003, **24**, (3), pp 27–29

[24] Luo, T -H : ‘Investigation on image threshold segmentation method of pests in stored grain’, *J. Wuhan Polytech. Univ.*, 2006, **25**, (1), pp 5–8

[25] Xinwen, Y , Zuorui, S : ‘Segmentation technology for digital image of insects’, *Trans. Chin. Soc. Agric. Eng.*, 2001, **17**, (3), pp 137–141

[26] Prakash, J , Vijayakumar, V : ‘A new texture based segmentation method to extract object from background’, *Glob. J. Comput. Sci. Technol. Graph. Vis.*, 2012, **12**, (15), pp 47–53

[27] Yuehua, C , Xiaoguang, H , Changli, Z : ‘Algorithm for segmentation of insect pest images from wheat leaves based on machine vision’, *Trans. Chin. Soc. Agric. Eng.*, 2007, **23**, (12), pp 187–191

[28] Zhao, J , Liu, M , Yao, M : ‘Study on image recognition of insect pest of sugarcane cotton aphis based on rough set and fuzzy c-means clustering’ IEEE Int Symp on Intelligent Information Technology Application, 2009, vol **2**, pp 553–555

[29] Yogamangalam, R , Karthikeyan, B : ‘Segmentation techniques comparison in image processing’, *Int. J. Eng. Technol.*, 2013, **5**, (1), pp 307–313

[30] Ramesh, R , Kulkarni, A C , Prasad, N R , *et al.*: ‘Face recognition using snakes algorithm and skin detection based face localization’ Int Conf on Signal, Networks, Computing, and Systems, 2016, pp 61–71

[31] Wang, Y , Peng, Y : ‘Application of watershed algorithm in image of food insects’, *J. Shandong Univ. Sci. Technol., Nat. Sci.*, 2007, **26**, (2), pp 79–82

[32] Zhang, W F , Guo, M : ‘Stored grain insect image segmentation method based on graph cuts’, *Sci. Technol. Eng.*, 2010, **7**, pp 1661–1664

[33] Mele, K : ‘Insect soup challenge: segmentation, counting, and simple classification’ IEEE Int Conf on Computer Vision Workshops, 2013, pp 168–171

[34] Yalcin, H : ‘Vision based automatic inspection of insects in pheromone traps’ IEEE Int Conf on Agro-Geoinformatics, 2015, pp 333–338

[35] Tibshirani, R , Walther, G , Hastie, T : ‘Estimating the number of clusters in a data set via the gap statistic’, *J. R. Stat. Soc., Ser. B (Stat. Methodol.)*, 2001, **63**, (2), pp 411–423

[36] Meyer, F : ‘Color image segmentation’ IET Int Conf on Image Processing and its Applications, 1992, pp 303–306

[37] Yao, Q , Liu, Q , Dietterich, T G , *et al.*: ‘Segmentation of touching insects based on optical flow and Ncuts’, *Biosyst. Eng.*, 2013, **114**, (2), pp 67–77

[38] Wang, Y , Weng, G : ‘The monitoring population density of pests based on edge-enhancing diffusion filtering and image processing’ Int Conf on Computer and Computing Technologies in Agriculture, 2007, pp 899–907

[39] Stehman, S V : ‘Selecting and interpreting measures of thematic classification accuracy’, *Remote Sens. Environ.*, 1997, **62**, (1), pp 77–89