A Cognitive Vision Approach to Early Pest Detection in Greenhouse Crops

Paul Boissard, Vincent Martin, Sabine Moisan

A Cognitive Vision Approach to Early Pest Detection in Greenhouse Crops

Paul Boissard [a,*], Vincent Martin [b], Sabine Moisan [b]

[a]*INRA-URIH, 400 route des Chappes, BP 167, 06903 Sophia Antipolis Cedex, France*

[b]*INRIA, ORION team, 2004 route des Lucioles, BP93, 06902 Sophia Antipolis Cedex, France*

Abstract

Early disease detection is a major challenge in horticulture. Integrated Pest Management (IPM) (Stern et al., 1959; Altner et al., 1977) combines prophylactic, biological and physical methods to fight bioagressors of crops while minimizing the use of pesticides. This approach is particularly promising in the context of ornamental crops in greenhouses because of the high level of control needed in such agrosystems. However, IPM requires frequent and precise observations of plants (mainly leaves) which are not compatible with production constraints.

Our goal is early detection of bioagressors. In this paper, we present a strategy based on advances in automatic interpretation of images applied to leaves of roses scanned *in situ*. We propose an cognitive vision system that combines image processing, learning and knowledge-based techniques. This system is illustrated with automatic detection and counting of a white fly (*Trialeurode Vaporariorum* Westwood) at mature stage. We have compared our approach with manual methods and our results showed that automatic processing is reliable. Special attention was paid to low infestation cases which are crucial to agronomic decisions.

Key words: IPM, greenhouse crops, pest early detection, cognitive vision, knowledge-based system, machine learning

* Corresponding author. Present address: INRA-URIH, 400 route des Chappes, BP 167, 06903 Sophia Antipolis Cedex, France. Tel.: +33 (0)4 92 96 26 57

  Email addresses: `Paul.Boissard@sophia.inra.fr` (Paul Boissard), `Vincent.R.Martin@sophia.inria.fr` (Vincent Martin), `Sabine.Moisan@sophia.inria.fr` (Sabine Moisan).

1 Introduction

A lot of research has been done on greenhouse agrosystems and more generally on protected crops to control pests and diseases by biological means instead of pesticides. Moreover, such agrosystems that are partly isolated from outside environment and highly controlled are good test sites for innovative methodologies in crop protection. A strong demand now exists in many countries for non-chemical control methods for pests or diseases, and this issue has not been studied enough (Hanafi, 2003). Greenhouses are considered as biophysical systems with inputs, outputs and control process loops. Most of these control loops are automatized (Ehret et al., 2001) (e.g., climate and fertirrigation control). However, no automatic methods are available to precisely and periodically evaluate the biotic status of plants. In fact, in production conditions, greenhouse staff periodically observe plants and search for pests. These observations may raise an alarm and point out places of dysfunction but they are qualitative and their accuracy depends on the human eye resolution, even if magnification tools can be used. This article takes place in a large-scale and multidisciplinary research program ultimately aiming at reducing pesticide use. To this end delivering pertinent information about the sanitary status of crops is essential because it is the starting point of decision making concerning crop protection and, moreover, it may allow to check whether a biological or chemical treatment has reached its objective.

In this paper, we focus on early pest detection. First, this implies to regularly observe the plants. In an automatic system this is done by digital remote sensors (e.g., imaging camera, scanner). Second, it is necessary to interpret image contents in order to identify objects corresponding to potential pests. Third, some actions (e.g., introduction of bio auxiliaries) must be undertaken if necessary. The focus of this paper is on the second point. Our goal is to automate *in situ* early detection of pests by a multidisciplinary approach.

1.1 Early detection

Early detection of symptoms (fungi) or of the initial presence of a bioagressor is a key-point in the context of IPM. Due to temperature and hygrometric conditions in greenhouses, rapid decisions are essential to control diseases and pests in order to avoid dissemination and permanent infestation (Van Lanteren and Woets, 1988; Lapchin and Shtienberg, 2002). (Malais and Ravensberg, 2003) give typical data for several pests about development time as a function of temperature and about female egg production. For instance, a white fly *Trialeurodes Vaporariorum* female on gerbera may lay about 130 eggs.

2

The detection of biological objects as small as such insects (dimensions are about 2 mm) is a real challenge, especially when considering greenhouses dimensions (10 to 100 meters long). Traditionally, visual observations are made each week by human experts, often on colored sticky traps. Since this technique does not allow to precisely study the epidemic spatial model, observations on natural support (i.e. on leaves) are preferred. But it is difficult or even not possible to perform a continuous (typically daily) control and to examine every leaf in the greenhouse. Adapted sampling could significantly contribute to reduce the amount of data and to speed up the analysis. Consequently, we need to reduce the 3D volume of canopy to be investigated. Investigated parts should be those firstly attacked or those with the greatest density of insects at an interesting development stage (e.g., growing stems for white fly larvae). We propose to sample the greenhouse by randomly cutting leaves in these parts.

1.2   Multidisciplinary strategy

Our objective is to develop a detection system that is robust and easy to adapt to different applications. We promote a multidisciplinary *cognitive vision* approach that integrates techniques from different domains, namely computer vision and artificial intelligence. Those domains can complement in two ways. First, artificial intelligence learning may help to fine tune vision algorithms so as to adapt them to various application needs. Second, the raw results of vision algorithms can be further processed by knowledge-based systems to check them, a way to improve robustness.

We turned to computer vision because even if radiometry has been used to quantify plant health deterioration (insect or fungi) (Xu et al., 2007; Mirik et al., 2006; Yang et al., 2005), it requires costly equipments complex to handle for non experts. Computer vision methods are easier to apply: in our system we simply use a consumer electronics scanner to get high resolution images of leaves. Moreover, radiometry is mainly applied to early detection of physiological stress due to a plant health attack (Delalieux et al., 2007; Chaerle and Van Der Straeten, 2000), which is not our purpose. Computer vision has a wider field of application such as disease and pest control. It has been applied in (Moya et al., 2005; Granitto et al., 2005; Skaloudova et al., 2006) respectively to quantify symptoms of powdery mildew, weed control and spider mite attacks, or in (Giacomelli et al., 1996) to develop an automated plant monitoring system in greenhouses. (Bauch and Rath, 2005) propose a complex vision-based system to measure white fly density within plant stands. But these computer vision systems are too specific to be easily re-usable and extended because their algorithms for image processing and interpretation are tightly tuned for their applications.

However, computer vision techniques are often not sufficient to obtain high quality detection, we propose to complement them with artificial intelligence methods. Hence, the image processing algorithms that we have developed are made auto-adaptive by learning techniques and completed by a knowledge-based approach. We obtain a rather robust system, where the weaknesses of a method are balanced by applying another one: image processing (IP) is made more versatile by learning methods and the raw numerical results of image processing can be interpreted at a higher conceptual level by knowledge-based techniques.

The organization of the paper is as follows: section 2 describes material and methods used for rose disease detection. Section 3 presents an overview of the proposed cognitive vision system, and details the system architecture. In section 4, we present how we apply this approach to white fly early detection on rose leaves. Results are presented and discussed in section 5. Section 6 summarizes our conclusion and outlines further research in this area.

2 Material and Methods Part 1: Constraints and Requirements related to the Greenhouse Agrosystem

As emphasized by other authors (Boll et al., 2007), two types of ratios are to be kept in mind when experimenting a method to detect bioagressors and to estimate their population and density. For instance, interesting indications may be given by the ratio of data relevance versus sampling duration (i.e. obtaining a set of significant data in a minimum time) and by the survey ratio (i.e. the foliar surface scanned versus the total foliar surface). To get realistic results, we have conducted an experiment in real conditions. Hence, we had to choose target crop and pest, a real greenhouse equipped with sensors, and a sampling strategy, that will be detailed in the following sub-sections.

2.1   Choosing a crop and a bioagressor as a model study

For this study, we first chose a model "crop × bioagressor". On the one hand, rose, an ornamental crop, was chosen because it attracts various bioagressors and it requires high level standard quality for flowers as well as leaves. On the other hand, white fly *Trialeurodes Vaporariorum* was chosen because this bioagressor requires early detection and treatment to prevent durable infestation. Eggs and larvae identification and counting by vision techniques are difficult because of critical dimension (eggs) and weak contrast between object and image background (larvae). For these reasons we decided to focus first on

4

adults. Eggs, larvae and adults are present on back faces of leaves. Since adults may fly away, we chose to scan the leaves when flies were not very active.

2.2 Greenhouse experiment

The agrosystem was a 256 m^2 plastic twin-tunnel greenhouse planted with roses and equipped with an opening roof, heating and a fog generator [1]. We used IPM for phytosanitary protection of plants. The management of climate, fertilization and irrigation was carried out by a control/command computer system designed at INRA. Two rose cultivars (Suella^{TM}, a yellow one, and Miss Paris^{TM}, a red one) were planted. They are known for their different resistance to powdery mildew and attractive powers to insects. Plants were grown in rock wool tubs of one meter length, ranged along six double lines at a density of 4.5 plants per m^2. Each line was divided into two parts: one half planted with yellow cultivars and the other one with red cultivars. The total cultivation corresponds to 1200 plants. A map of the greenhouse is shown on Figure 1.
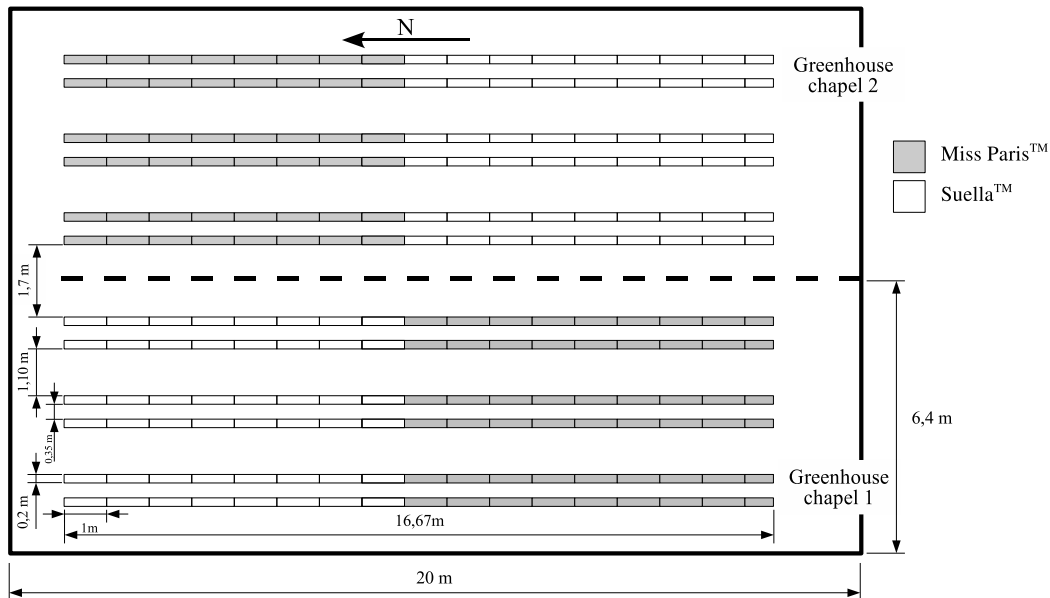

Fig. 1. Greenhouse map showing two chapels of 128 m^2 each.

2.3 Sampling strategy

We chose our sampling strategy based on the following requirements:

---

[1] average size of droplets: 19 μm

- Spatially, data should be uniformly distributed, thus samples were randomized (according to a grid) over the whole greenhouse.
- Temporal sampling should be realistic, i.e. provide a good ratio data relevance versus sampling duration.

The spatial sampling strategy consists in a randomized sampling in the horizontal plane and optimized sampling along the vertical axis. Since it is convenient to consider 12 plants or 2 m rock wool slab as a standard visual observation unit, it was decided to take a leaf sample (5 or 7 leaflets) every second meter along plantation lines. We have done a pre-study on sample cuts at various heights of plant canopy to decide the optimal localization of samples: for early detection of mature white flies, growing stems have been preferred. Hence, 100 samples were taken corresponding to 1200 plants. Samples are rose leaves, each leaf being made up of 5 or 7 leaflets, cut in the central part of the canopy where growing stems are the most numerous. Both sides of leaves were scanned individually and 200 images were recorded (see Figure 2 for an example). If we assume a LAI (Leaf Area Index) of 3 for rose crop (Raviv and Blom, 2001), and with an effective crop surface of 100 m^2, it means that around 0.36% of LAI (for one face) is analyzed at each survey by using the above sampling strategy[2].

Concerning the temporal sampling strategy, the time required to perform an automatic survey is of the same order of magnitude as the time necessary to make a chemical treatment on an equivalent surface. Thus, this quick delivery of results, i.e. within half a day, is compatible with rapid decision.
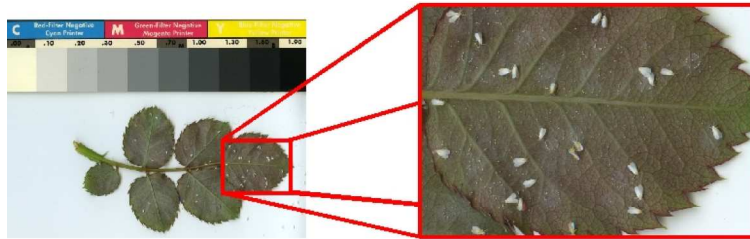

Fig. 2. Example of a scanned rose leaf infested by white flies.

2.4 Sensor and calibration

For this study, samples were manually cut and scanned directly in the greenhouse by using a consumer electronics flatbed A4 scanner[3]. The scanner was

---

[2] The survey ratio (SR) is computed as follows: $SR = (N_{scan} * S_{scan})/(LAI * S_{crop})$, with N_{scan} the number of total scanned leaves (200), S_{scan} the effective scanned surface per acquisition (0.0054 m^2) and S_{crop} the effective crop surface (100 m^2)

[3] *Epson Perfection 4990 photo*

connected to a laptop via a USB port. This allowed a high image quality and a short scanning/transfer time. A resolution of 600 dpi was chosen. This corresponds to theoretical square pixel dimensions of 42 μm × 42 μm. Such a resolution is a good compromise: it is precise enough to digitize objects as little as mature white flies (500 pixels of area) and compatible with data acquisition/storage constraints. Leaves are scanned individually together with a reference colored chart to calibrate the scanner. Each color image of 3000x4000 pixels is recorded in PNG format corresponding to about 20 mega octets of data. The 200 scanned leaflets represents a total of 4 gigaoctets of data.

Once the data acquisition conditions fixed, the next step is to provide a system that automatically identifies and counts white flies on the scanned images. This system is presented in the next section.

3 Material and Methods Part 2: Cognitive Vision System

Traditional manual counting is tedious, time consuming and subjective, for it depends on observer's skill. To overcome these difficulties, we propose to automate identification and counting, based on computer vision and artificial intelligence techniques.

During the last three decades computer vision systems have been widely used to automate image understanding. Application domains are as various as medicine, biology, agronomy, or surveillance. But in most cases, for convenience reasons, the general vision problem and the application one are strongly interrelated. This leads to build *ad hoc* systems which are difficult to re-use for other applications.

Artificial intelligence knowledge-based approaches, relying on explicit knowledge, enable a clear separation between the different underlying problems of vision systems (i.e. mainly image processing and high-level interpretation). A good overview of knowledge-based vision systems can be found in (Crevier and Lepage, 1997). However, two open problems still remain: first, knowledge acquisition bottleneck when a large amount of knowledge is needed and, second, lack of robustness when faced with varying conditions.

To overcome the brittleness of classical vision systems, a new discipline called cognitive vision has recently emerged. A research road-map has been proposed by a European Network (ECVISION, 2005) and a tentative framework for this discipline can be found in (Vernon, 2006). Cognitive vision proposes to enrich computer vision systems with *cognitive* capabilities, e.g., to reason from *a priori* knowledge, to learn from perceptual information, or to adapt its strategy to different problems. This new discipline thus involves several existing

related ones (computer vision, pattern recognition, artificial intelligence, cognitive science, etc.). Some systems have started to implement cognitive vision ideas, mainly for human behavior recognition, relying on different technologies, especially in learning. For example, in (Vincze et al., 2006) a cognitive system combining low-level image components and high-level activity reasoning processes has been developed to recognize human activities. This system integrates various techniques such as connectionism, Bayesian networks, component framework, and robotics. Another example (Nagel, 2004) converts a video sequence into a natural language description of "what is going on" in road traffic scenes. It uses a Fuzzy Metric-Temporal (Horn) Logic (FMTHL) and *a priori* knowledge of 3D models of cars. Contrary to our approach, some authors (Granlund, 2005) promote a fully autonomous system without the need of any *a priori* knowledge. All the necessary information should be learnt for instance by neural learning techniques.

In the line of the cognitive vision approach, we propose an automatic image interpretation system that combines image processing, neural learning and knowledge-based techniques. This approach follows our previous work presented in (Hudelot and Thonnat, 2003) and (Boissard et al., 2003).

The necessary information to build such a system comes from different domains. Descriptive symbolic knowledge about insects exists in public biological data bases[4]. Image processing programs (e.g., segmentation algorithms) have been widely experimented and are even available in popular libraries. The contribution of a cognitive vision approach is to integrate and to structure all this information with the help of learning and knowledge representation techniques.

3.1 System Overview

As human biologists do, our system has to analyze raw images and to label interesting regions that correspond to objects of interest (e.g., insects). To recognize a region as an insect, a human expert relies on (biological and contextual) domain knowledge about insects (e.g., species features, life cycle, host plant) as well as visual data that can be extracted from images (color, texture, shape, and size). A software system must take into account both kinds of knowledge. To separate the different types of knowledge and the different reasoning strategies, we propose an architecture based on specialized modules, as shown in Figure 3. It consists of two knowledge-based systems (KBS), a set of image processing (IP) algorithms, and an initial learning module.

Before routine execution, a *learning* stage (Figure 3 top) is performed once on

---

[4] e.g., see white fly description at **http://www.inra.fr/hyppz/glossary.htm**
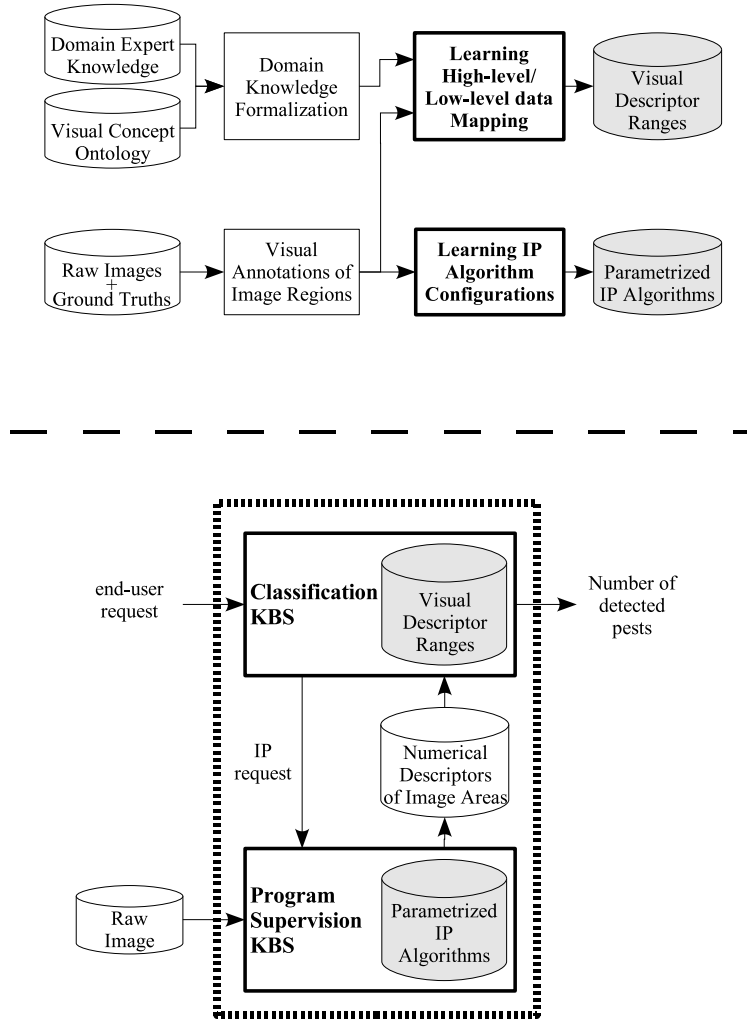
Fig. 3. Cognitive vision system. The top part corresponds to the initial learning module and the bottom part to the automatic system for routine execution.

a training image sub-set. This preliminary stage is used to complement the knowledge necessary to run the two following KBSs.

The *classification* KBS (Figure 3 bottom) aims at selecting interesting regions in images. To this end it triggers image processing requests and interprets the numerical results into higher level concepts, i.e. (parts of) objects of interest. It retains only the regions corresponding to target insects and returns their number to the user.

The *supervision* KBS (Figure 3 bottom) is used to monitor the execution of image processing requests. It selects and plans the best programs with the best parameter values for each image. From raw images provided by the end-user, the goal is to extract numerical values needed by the classification KBS.

The following sub-sections detail these three different parts.

9

3.2   Learning Stage

Learning techniques are used for two purposes: to learn how to map low-level features to high-level domain concepts and how to tune parameters of image processing programs.

3.2.1   Learning Visual Concepts

The goal is to map object descriptions in the knowledge base of the classification system to numerical values. Most real objects can be described in terms of concepts, such as their shape, color or texture. We call them *visual concepts*. Visual concepts are an intermediate level that helps mapping low-level numerical value to a domain class description. For instance, a length in millimeters and/or a color in RGB values may be mapped to an insect body. Thus, each biological class description appearing in the classification knowledge base must be precisely specified in terms of visual concepts.

We refer to a general ontology proposed by N. Maillot et al. in (Maillot and Thonnat, 2006), which is a hierarchy divided into three parts: spatio-temporal, color and texture concepts. For instance, spatio-temporal concepts include shape, size, and spatio-temporal relations. The main advantage in using this ontology is to provide domain experts with a vocabulary for describing domain classes in visual terms (as shown in Figure 5) by means of *numerical descriptors*. The role of these descriptors is to bridge the semantic gap between low-level numerical values and visual concepts. A descriptor has an attached numerical value that can be computed by vision programs; for instance, a program can compute a region length in millimeters. It also corresponds to a visual feature meaningful for a human expert to describe an object; for instance, an expert may select the *length* descriptor of the *size* concept as relevant to describe an insect body.

Based on a training set of images and for each visual concept used by the expert, the learning module learns the range of possible values for all the numerical descriptors necessary to recognize the concept.

3.2.2   Learning Image Processing Parameters

In our approach, we paid a special attention to image segmentation. Segmentation consists in grouping pixels sharing some common characteristics (e.g., color distribution, topology or texture). But defining a criterion for grouping pixels clearly depends on the segmentation goal (e.g., segment insects, segment leaf veins, etc.). Consequently, a unique general method cannot fit all goals. When designing a vision system, segmentation algorithms are often

heuristically selected and narrowly tuned by an image processing expert with respect to the application needs. Generally, such a method leads to *ad hoc* algorithms working under fixed hypotheses, in particular with fixed values of free parameters. Although default parameter values are provided by authors of the algorithms, these parameters need to be tuned to get meaningful results (see an example in Figure 4).

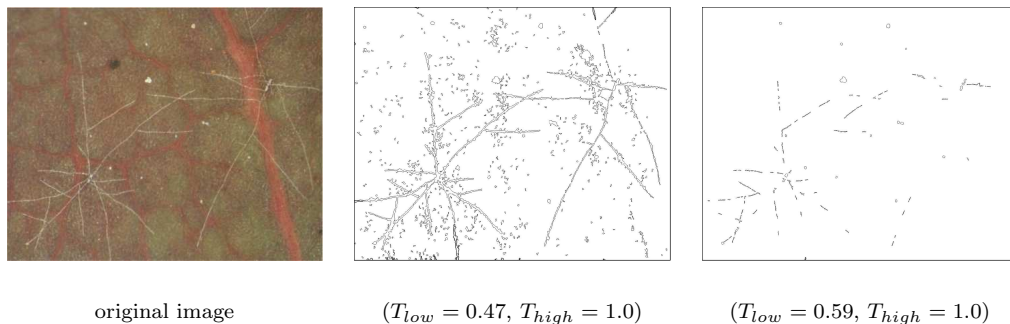| original image | $(T_{low} = 0.47,\ T_{high} = 1.0)$ | $(T_{low} = 0.59,\ T_{high} = 1.0)$ |

Fig. 4. Illustration of the influence of parameter setting: a segmentation algorithm, Hysteresis thresholding, is tuned with two different values for its two control threshold parameters (T_{low}, T_{high}). Correct parameter values might be between these.

As detailed in (Martin and Thonnat, 2007), we propose to use learning techniques for adaptive image segmentation. No new algorithms are proposed, but rather a methodology that allows to easily set up a segmentation algorithm in a vision application. More precisely, we propose a learning approach for parameter tuning according to the image contents and the application needs. This supervised off-line learning approach relies on hand-labeled samples that corresponds to different possible contexts. We call *context* a set of discriminant visual characteristics, shared by several images in the training set. Contexts may correspond to different camerawork conditions or application-dependent situations. For instance, in our application rose leaves can take different aspects (e.g., color, density and contrast of leaf veins) according to their maturity or to the observed side. The different contexts are identified by clustering techniques on training images. Considering that white flies are semi-transparent, it is clear that segmentation cannot be configured with a fixed set of parameters (e.g., thresholds) for all contexts. Instead, our approach learns sets of acceptable values for each algorithm parameter and for each context from a set of images sharing common characteristics. This allows to adapt the parameter setting to the different image contexts identified in the training set.

3.2.3 Learning Issues

At the end of the learning stage, we get image processing algorithm with fine tuned parameters and descriptor ranges for all relevant numerical descriptors of the application.

11

Note that this learning stage is done only once, for one (set of) insect(s) to detect and one set of programs to run. Provided that the acquisition conditions do not change drastically, routine execution only involves running the two following knowledge-based systems.

3.3 Classification System

The role of the classification system is to recognize and to count white flies on an image. To this end its relies on knowledge about insect descriptions and on numerical descriptor values provided by image processing programs.

The knowledge in this KBS consists of descriptions of domain classes and class hierarchies, provided by domain experts. We propose a dedicated expert language to describe these hierarchies. In the case of white fly detection, the knowledge base mainly contains knowledge about these insects (see Figure 5). The *WhiteFly* domain class describes how such an insect may be recognized thanks to different visual concepts selected by the expert, namely its shape, size and color. These concepts refer to general ones (such as *ColorConcept*) defined in the general ontology.

DOMAIN CLASS *WhiteFly* SUPERCLASS: *Bioagressor*					
SHAPE: ShapeConcept					
Descriptors:					
circularity	[	0.05	0.2	0.5	0.6	]
excentricity	[	0.1	0.2	0.4	0.5	]
rectangularity	[	0.5	0.6	0.8	0.85	]
elongation	[	0.3	0.35	0.7	0.8	]
convexity	[	0.7	0.75	1.0	1.1	]
compacity	[	0.1	0.25	0.9	1.0	]
COLOR: ColorConcept					
Descriptors:					
saturation	[	0.0	0.0	0.2	0.3	]
lightness	[	120	130	240	260	]
hue	[	80	90	170	180	]
SIZE: SizeConcept					
Descriptors:					
area	[	0.5	0.6	1.2	1.3	]
length	[	0.6	0.8	2.5	3.5	]
width	[	0.2	0.3	1.0	1.3	]

Fig. 5. High-level description of a domain class (white fly). Visual concepts are in SMALL CAPS. Learned fuzzy ranges are shown on the right. They are composed of four numbers, corresponding respectively to the minimum admissible value, the minimum and maximum most probable values, and the maximum admissible value.

Each visual concept is in turn described by a relevant set of numerical de-

scriptors and their associated fuzzy ranges of possible values, as learned on the training set with the help of a domain expert during the initial learning stage (see section 3.2.1). For example, some values of the HSV color of image areas are linked with the white fly color. Experts use the vocabulary defined by the general visual concept ontology such as the term "circularity" to characterize the shape of a white fly.

It should be noted that we do not need to manage a complete biological hierarchy of insects (i.e. with all sub-species), but only the parts that are useful for the recognition task. Indeed, it is useless (and often impossible with the currently available vision techniques) to precisely recognize the sub-species of an insect, because we know that not all sub-species will infest a type of plant.

To summarize, the classification KBS provides class hierarchies and a description of each class in terms of numerical descriptors. To get the real values of numerical descriptors, the correct image processing algorithms must be launched and controlled: this is the role of the image processing supervision KBS.

3.4 Image Processing Supervision System

The image processing task itself is achieved by a program supervision (Thonnat and Moisan, 2000) knowledge-based system. Program supervision techniques make it possible to automate the use of complex programs, i.e. to plan and control processing activities. In cognitive vision systems, the supervision system controls the use of the image processing sub-tasks, such as image segmentation and features extraction. It is based on knowledge about the programs, their input/output data, their applicability conditions, their possible combinations, and the necessary information to run them in various situations. This knowledge is given by image processing experts in the form of *operators* and decision *rules* to guide the supervision engine reasoning (e.g., to select programs, initialize their parameters, or assess their results). Suitable parameter initialization values have been learned during the previous learning step (see 3.2.2), but more tuning is possible dynamically depending on the input image specificity.

Operators and rules are formally described in the knowledge base using an expert language (see an example Figure 6). Figure 7 shows a sketchy view of the current organization of operators. It corresponds to different possible plans, among which the system will choose the best path, i.e. the plan of programs which is best suited for each image.

Let us illustrate the processing of an image through the supervision system. Given an initial image (such as the one in Figure 8 left), the knowledge-

```
Composite Operator { name Segmentation
     comment ''image segmentation operator''
     Input Data
          Image input_image
     Output Data
          Image output_image
          ...
     Body
          RegionBased | EdgeBased
     Choice Rule { name RegionChoice
          If concept == ShapeConcept
          Then useOperator RegionBased }
}
```

Fig. 6. An example from the program supervision knowledge base. A composite operator describes an alternative decomposition (denoted by a |) into two sub-operators: region or edge-based segmentation, and a rule selects the first one if the concept to recognize (as indicated by the classification KBS) is Shape.
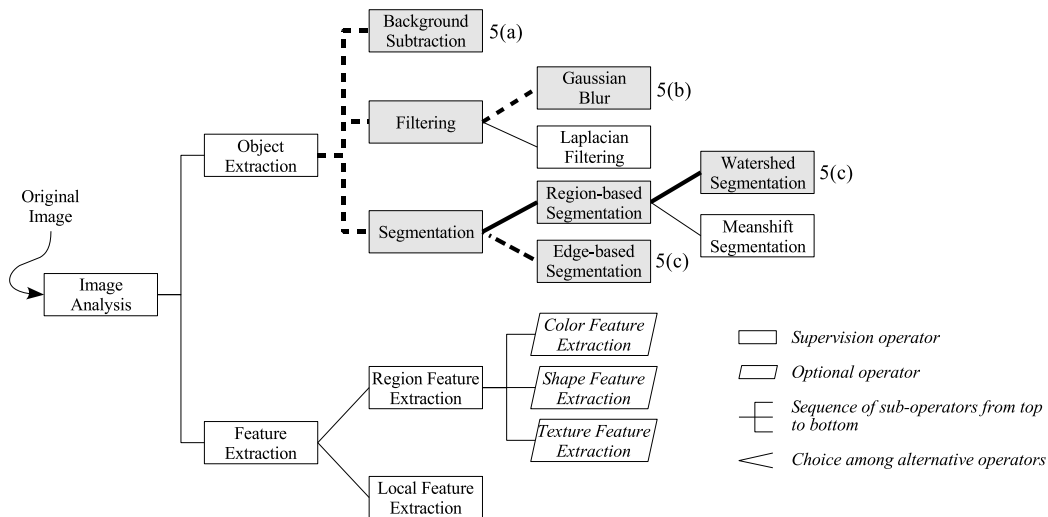

Fig. 7. Graph of the main image processing operators in the supervision knowledge base.

based system launches the first operator, *Image Analysis*, which breaks up into a sequence of two sub-operators, as shown in Figure 7: *Object Extraction* followed by *Feature Extraction*. *Object Extraction* itself decomposes into a sequence (*Background Substraction*, then *Filtering*, and finally *Segmentation*). Since *Background Substraction* appears on the top and corresponds to a concrete program to execute, the system invokes it. This program automatically extracts a leaf from its background image (Figure 8(a)). The second sub-operator, *Filtering* may be performed in two different ways (Gaussian or Laplacian filtering), one of them is automatically chosen by the KBS, using expert decision rules. In our case *Gaussian Blur* is chosen. It runs the corre-

14

sponding denoising program and the result is presented on Figure 8(b). The next operator, *Segmentation*, also corresponds to a choice between two alternative sub-operators: *Region-Based* or *Edge-Based*. Using expert choice rules in the knowledge base (such as the one presented in Figure 6), the KBS selects *Region-Based*, then *Watershed Segmentation* to segment the image into regions. The result after segmentation is shown in Figure 8(c). The system has thus applied the following plan of programs for the extraction step: *Background Substraction*, *Gaussian Blur* and *Watershed Segmentation*. It corresponds to the bold path in Figure 7. Note that for another image, such as the powdery mildew one in Figure 9 left, the best plan is different. Thanks to its knowledge base, the system will select *Background Substraction*, *Gaussian Blur* and *Edge-based Segmentation* (dashed path in Figure 7).
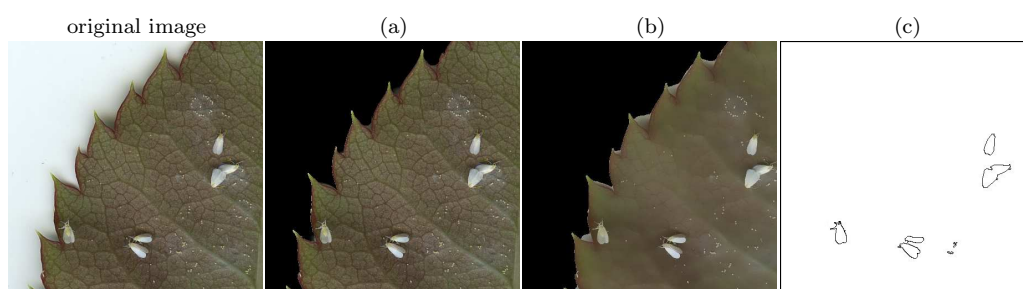

Fig. 8. Image processing flow for white fly extraction. From left to right: input image, (a) background subtraction, (b) filtering result, (c) segmentation result.
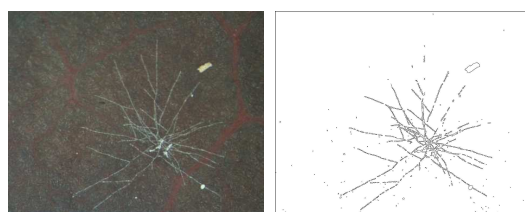

Fig. 9. Image processing for powdery mildew extraction.

Similarly, once the objects extracted, the second step of *Image Analysis*, feature extraction, computes the attributes corresponding to each region, according to the domain feature concepts (e.g., color, shape and size descriptors) and to the operator graph. The process runs up to the last program in the decomposition (in the example, it appears to be *Shape Feature Extraction*). Finally, the supervision KBS returns a set of candidate areas to the classification KBS. Each area has an attached set of computed numerical descriptor values. The classification KBS can use these descriptors in conjunction with its own knowledge to select areas corresponding to white flies and to return the number of recognized flies to the user.

4 Results

The system has been tested on a database composed of a representative sample of 200 images of scanned rose leaves (see section 2.2) provided by INRA-URIH (Sophia Antipolis, France). At that time the greenhouse was infested with white flies, mainly localized at the lower face of leaves. The study target pest was mature white flies and the purpose of the system was to identify and count them. The ranges of values of corresponding numerical descriptors (e.g., size, color, shape) have been semi-automatically learned on a subset of 20 images. The global computing time for an image of 2495×4056 pixels is around 35 seconds with a 2,33 Ghz Dual Core Xeon processor.

To assess the quality of our cognitive system, the results have been compared with ground truth. Three human operators (one expert in agronomy, one expert in image processing and one non-expert neither in agronomy nor in image processing) have manually counted the white flies on 180 images. Each operator has a different point of view when counting. The expert in image processing focuses on pure visual characteristics while the expert in agronomy focuses more on the semantic meaning of images. This can lead to different counting results as illustrated in Figure 10: the expert in agronomy counts three white flies, the expert in image processing only one (because only one object matches the visual criteria), and the non-expert two. This explains the size of some error bars on ground truths in Figure 11 (e.g., sample 142 and 148).

Fig. 10. Example of an ambiguous image sample for ground truth estimation. The two white flies on the top have moved during the scanning. This leads to color flickering which does not correspond to the normal white fly color.

We detail hereafter the result evaluation for early detection of mature white flies. From the 180 images composing the test set, 162 contain between zero and five white flies. Figure 11 presents the detailed results for the whole test set. For each image the average ground truth value (blue circle) is reported with its associated error bar. Red crosses represent the values found by the system. To prove the reliability of our learning approach, we have tested it against an *ad hoc* segmentation (i.e. a manually tuned algorithm): a hysteresis thresholding segmentation on gray-scaled normalized image (i.e. pixel values in $[0, 1]$) with low threshold (T_{low}) fixed to 0.45 and high threshold (T_{high}) fixed to 1.0. The two graphs present the results of mature white fly counting. The

first graph corresponds to the system configured with *ad hoc* segmentation and the second one corresponds to the system configured with our learning approach.
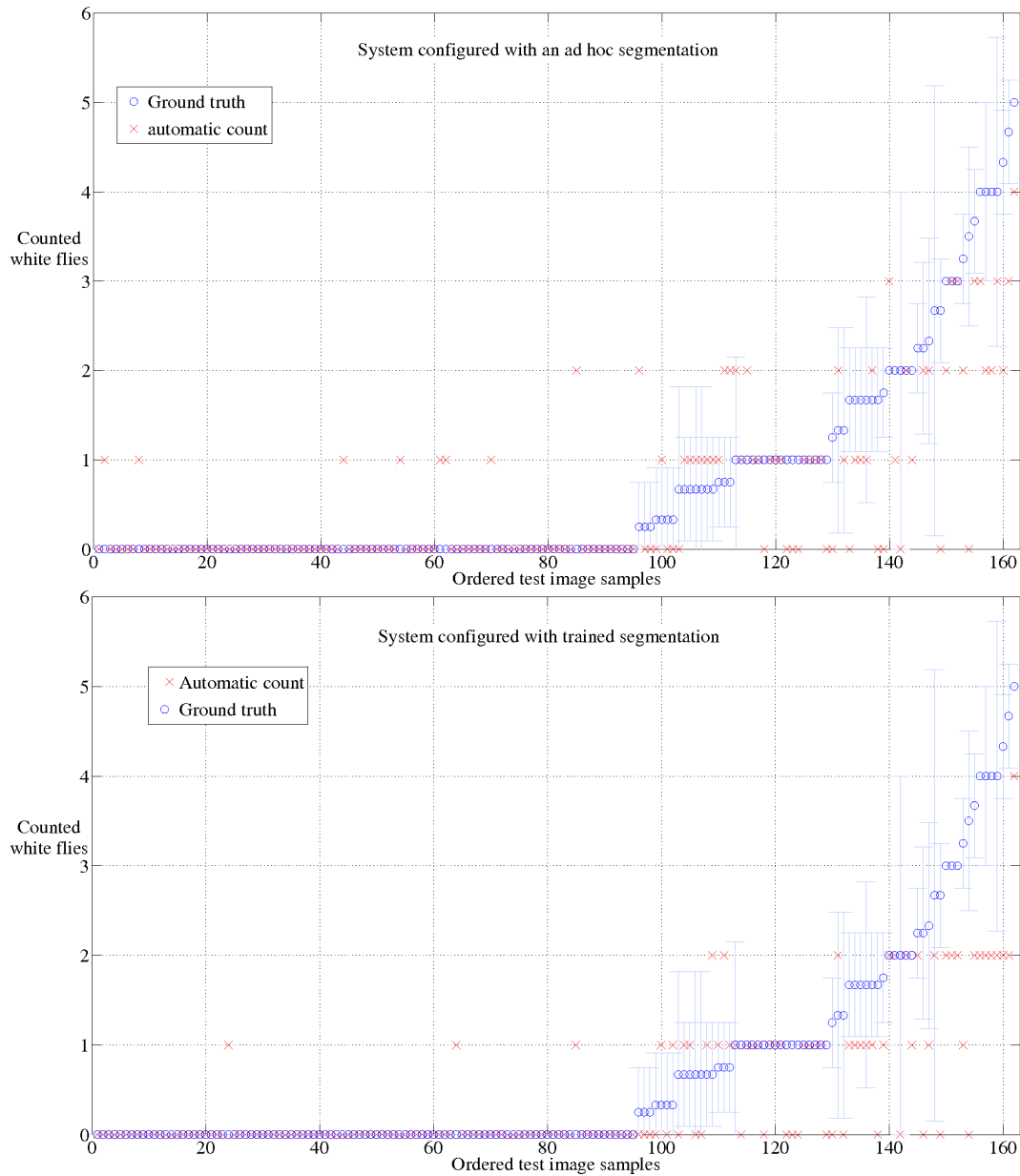

Fig. 11. Evaluation of mature white fly counting results in early detection cases (i.e. between 0 and 5 flies per leaf). The upper one presents the results for the system configured with an *ad hoc* segmentation, the lower graph presents the results for the system configured with trained segmentation parameters.

Globally, the detection rate is satisfactory. To fully make use of the results, we can separate the test samples into two classes depicting the most relevant situations. The first class (C_1) represents images with no mature white fly at all (i.e. images for which the maximum of ground truth error bar is strictly

inferior to 1.0) and the second class (C_2) represents images with at least one white fly detected (i.e. images for which the maximum of ground truth error bar is equal or superior to 1.0). We define the False Positive Rate (FPR) as the rate of over-detection (i.e. images for which the number of detected white flies is greater than the ground truth error bar) and the False Negative Rate (FNR) as the rate of under-detection (i.e. images for which the number of detected white flies is less than the ground truth error bar). Table 1 summarizes the detection results. The figures represent the mean values of FNR and FPR for class C_1, C_2, and for the whole image test set.

Table 1
False Negative Rate (FNR) and False Positive Rate (FPR) for test images with no white flies (class C_1), at least one white fly (class C_2) and for the whole test set.

| | Results for early detection of mature white flies | | | |
| | With ad hoc segmentation | | With learned segmentation | |
Samples	FNR (%)	FPR (%)	FNR (%)	FPR (%)
C_1 (102)	-	9.6	-	3.1
C_2 (60)	24.7	9.0	29.6	2.0
Whole set (162)	9.1	9.4	11.0	2.7

5 Discussion

The FNRs are roughly similar for the two configurations. In fact, this reveals confusing situations as the one presented in Figure 12: two overlapping white flies have been segmented into one region which has obviously not the shape of a single white fly. Hence, the system counts one white fly instead of three. This highlights the scale issue of our problem for which highly variable small objects are expected to be detected in a complex natural environment. Concerning the FPRs, they are up to four times smaller with the learned segmentation than with the *ad hoc* one. This is due to our adaptive segmentation approach that allows to efficiently tune algorithm parameters with respect to variations in leaf color and contrast.
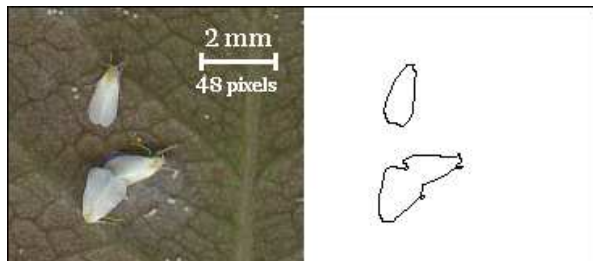

Fig. 12. Example of an ambiguous situation leading to a wrong interpretation result.

We have also compared our results to a manual method of identification and counting [5] as shown in Table 2. This table raises several remarks. First, in the automatic mode, our prototype currently detects only one class of pest. New developments will allow us to identify more classes, in almost the same amount of time. Second, manual mode operations are achieved in a research laboratory with optical magnifying tools and no time constraints; however under production conditions, rapid observations are necessary. Third, the manual mode is limited by human capacities about objects counting and discrimination ability which progressively decrease while the time passes due to tiredness and repetitivity.

Table 2

Method comparison for identification and counting of pests in a 256 m^2 greenhouse.

	Manual identification and counting	Automatic identification and counting
Number of samples	240 randomized leaflets	100 leaves composed of 5 to 7 leaflets ×2 (faces)
Total cumulated time needed for all operations	from 3h30' (few pests) to 5h30' (lot of pests)	3h (sampling + scanning + storage) + 1h40' (identification + counting) = 4h40'
Average process time per sample	**52" to 1'22"** per leaflet	**12" to 17"** per leaflet
Identified classes (number of sub-classes)	**16**: white flies (4), enemies of white flies (2), trips (3), spider mites (3), beneficial organism (1), fungi (2), aphids (1)	**1**: white flies (1)
Fraction of LAI analyzed, with an assumed LAI of rose crop equal to 3	about **0,17%**	about **0,36%**
Delivery of results	**2 days**	**4h40'**
Advantages	Discrimination capacity	Accuracy not depending on time spent, quick delivery of result
Disadvantages	Need of a specialized operator (taxonomist) and precision vs time	In the current prototype, limited number of classes

Compared to the system proposed by Bauch and Rath (Bauch and Rath, 2005), our system appears simpler and more precise. Indeed, their system

---

[5] Personal communication from R. Brun, C. Métay-Merrien and M.C. Grosnier, INRA, 2007

uses an heavy experimental protocol combining a mobile suction mechanism to collect white flies on tomato crop with an optical recognition system to analyze image data. They concluded that their system is able to differentiate between high and medium infestation densities (i.e. respectively 75 and 150 white flies per sample) with a good recognition rate for correctly classified white flies (83%). However, results on low density measures are not very convincing since the high false classification rate leads to erroneous pest density quantification. This is a major problem because differentiation between no infestation class and medium infestation class is crucial for rapid decision making.

6   Conclusion

We propose an original approach for *in situ* early detection of bioagressors, which has been applied to detect mature white flies on rose leaves. To detect biological objects on a complex background, we combined scanner image acquisition, sampling optimization, and advanced cognitive vision. Our system is composed of extensible knowledge-based systems for image analysis and natural object recognition, coupled with image processing programs and machine learning. It illustrates the collaboration of complementary disciplines and techniques which led to an automated, robust and versatile system. The prototype system proved reliable for rapid detection of white flies. It is rather simple to use and exhibits the same performance level as a classical manual approach. Moreover, it detects white flies three times faster and it covers three times more leaf surface.

The context of our work is to automate operations in greenhouses. An active line of research concerns robotic applications (Noordam et al., 2005; Belforte et al., 2006), mainly for automating cutting and harvest. Our goal is rather to better spot the starting points of bioagressors attacks and to count these latter.

As previously mentionned in section 1, vision techniques have been widely used in agriculture, but most work led to application specific methods or systems, whereas we aim at proposing a *generic* approach, based on cognitive vision. There exist some works concerned with genericity, at least to some extend, or with cognitive properties. For instance, we can cite (Bernier and Landry, 2000) who have also as objective early disease detection, and who propose a versatile algorithm to identify and count fungal spores. They have also propose a generic method (Bernier and Landry, 2003) to represent and compare irregular natural shapes on images (leaf shapes). (Perner et al., 2006) introduce case-based reasoning to accomodate variations in the appearance of natural objects; the contents of their cases plays a similar role as our visual descriptions. (Manh et al., 2001) propose to introduce *a priori* knowledge about the shape of

objects to detect (in their case weed leaves) to guide the segmentation. In (El-Helly et al., 2004), the authors describe a biological expert system enhanced with vision techniques, while in our system it is rather the vision system which is enhanced by knowledge-based facilities. Closer to our approach, an experimentation framework, named Motris, is proposed in (Dahlkamp et al., 2007). This framework for 3D-model based tracking is flexible and modular to compare and combine algorithms; in our case this is the role of the program supervision system.

7 Future Work

The results presented in this paper are promising but several improvements in both material and methods can be carried out to reach the requirements of an IPM system.

Our first objective is to detect other white fly stages (eggs, larvae) and other bioagressors (aphids) or plant diseases (powdery mildew). Thanks to our cognitive approach, it is simple to introduce new objects to detect or new image processing programs to extract the corresponding information. Since knowledge is expressed in a declarative way, it is easier to modify than if deeped in code. Our approach is independent of the experimental protocol (sampling strategy or model study) instead we propose a generic architecture that can be adapted to different applications. The generic part of the system is re-usable, only parts of knowledge bases and image processing programs to call will change. Introducing new bioagressors will also require a new learning stage on a training set of images containing these insects, but afterward there will be only a low extra cost in execution time.

The second objective is to improve the camerawork in the greenhouse. While recognition is fully automated, image acquisition still requires manual operations for cutting and scanning leaves and problems remain to rapidly acquire images *in situ*. Moreover, scanners are sensitive to motion which may cause trouble as shown in Figure 10. We plan to overcome such limitations by using video cameras or other sensors (Wang et al., 2006). Another advantage of video cameras is that they provide temporal information which is of great interest to disambiguate occlusion situations as seen in Figure 12. This implies to integrate new algorithms for temporal aspects (e.g. motion detection and tracking). Continuous image acquisition will also allow us to investigate online incremental learning (Ross et al., 2007) of visual descriptors and of algorithm parameters. It seems also interesting to introduce deformable template models (Manh et al., 2001) at the segmentation level because it may improve difficult detection cases, not fixed by region-based segmentation.

Acknowledgments

References

Altner, G., Baggiolini, M.and Celli, G., Schneider, F., Steiner, H., 1977. La protection intégrée, une technique d'appoint conduisant à la production intégrée. Bulletin OILB/SROP 4, 117–132.

Bauch, C., Rath, T., 2005. Prototype of a Vision Based System for Measurements of White Fly Infestation. In: Acta Horticulturae (ISHS) 691. pp. 773–780.

Belforte, G., Deboli, R., Gay, P.and Piccarolo, P., Ricauda Aimonino, D., November 2006. Robot Design and Testing for Greenhouse Applications. Biosystems Engineering 95 (3), 309–321.

Bernier, T., Landry, J.A., 2000. Algorithmic Recognition of Biological Objects. Canadian Agricultural Engineering 42 (2), 101–109

Bernier, T.; Landry, J.A. 2003. A new method for representing and matching shapes of natural objects. Pattern Recognition 36 (8), 1711-1723.

Boissard, P., Hudelot, C., Thonnat, M., Peŕez, G., Pyrrha, P., Bertaux, F., September 2003. An automated approach to monitoring the sanitary status and to detect early biological attacks on plants in greenhouses - example on flower crops. In: International Symposium on Greenhouse Tomato, Integrated Crop Protection and Organic Production. Avignon, France.

Boll, R., Marchal, C., Poncet, C., Lapchin, L., 2007. Rapid Visual Estimates of Thrips (Thysanoptera: Thripidae) Densities on Cucumber and Rose Crops. Journal of Economic Entomology 100 (1), 225–232.

Chaerle, L., Van Der Straeten, D., November 2000. Imaging techniques and the early detection of plant stress. Trends in Plant Science 5 (11), 495–501.

Crevier, D., Lepage, R., 1997. Knowledge-based image understanding systems: a survey. Computer Vision Image Understanding 67 (2), 160–185.

Dahlkamp, H., Nagel, H.-H., Ottlik, A., Reuter, P., June 2007. A Framework for Model-Based Tracking Experiments in Image Sequences. International Journal of Computer Vision 73 (2), 139–157.

Delalieux, S., Van Aardt, J., Keulemans, W., Schrevens, E., Coppin, P., 2007. Detection of biotic stress (Venturia inaequalis) in apple trees using hyperspectral data: Non-parametric statistical approaches and physiological implications. European Journal of Agronomy, in Press.

ECVISION, August 2005. A Research Roadmap of Cognitive Vision. Tech. rep., IST Project IST-2001-35454.

Ehret, D. L., Lau, A., Bittman, S., Lin, W., Shelford, T., 2001. Automated monitoring of greenhouse crops. Agronomie: Agric. Environ. 21 (4), 403–414.

El-Helly, M., El-Beltagy, S., Rafea, A., 2004. Image analysis based interface for diagnostic expert systems. In: WISICT '04: Proceedings of the Winter International Symposium on Information and Communication Technologies. Cancun, Mexico, pp. 1–6.

Giacomelli, G., Ling, P., Morden, R., 1996. An Automated Plant Monitoring System Using Machine Vision. Acta Horticulturae (ISHS) 440, 377–382.

Granitto, P. M., Verdes, P. F., Ceccatto, H. A., 2005. Large Scale Investigation of Weed Seeds Identification by Machine Vision Techniques. Computers and Electronics in Agriculture 47 (1), 15–24.

Granlund, G., 2005. A Cognitive Vision Architecture Integrating Neural Networks with Symbolic Processing. Künstliche Intelligenz 2, 18–24

Hanafi, A., 2003. Integrated Production and Protection Today and in the Future in Greenhouse Crops in the Mediterranean Region. Acta Horticulturae (ISHS) 614, 755–765.

Hudelot, C., Thonnat, M., 2003. A cognitive Vision Platform for Automatic Recognition of Natural Complex Objects. In: IEEE (Ed.), International Conference on Tools with Artificial Intelligence, ICTAI, Sacramento. pp. 398–405.

Lapchin, L., Shtienberg, D., 2002. Integrated Pest and Disease Management in Greenhouse Crops. Vol. 14. Kluwer academic publishers, Dordrecht, Chap. Sampling and Monitoring Pests and Diseases, pp. 82–96.

Maillot, N., Thonnat, M., 2006. Ontology Based Complex Object Learning and Recognition. Image and Vision Computing Journal, Special Issue on Cognitive Computer Vision, in Press.

Malais, M., Ravensberg, W., 2003. Knowing and Recognizing. The biology of glasshouse pests and their natural enemies. Koppert B.V., The Netherlands.

Manh, AG., Rabatel, G., Assemat, L., Aldon, MJ., 2001. Weed leaf image segmentation by deformable templates. Journal of Agricultural Engineering Research 80 (2), 139–146.

Martin, V., Thonnat, M., 2007. Scene Reconstruction, Pose Estimation and Tracking., June, 2007, edited by Rustam Stolkin, Chap. A Learning Approach for Adaptive Image Segmentation, pp. 431–454.

Mirik, Jr., M., Michels, G., Kassymzhanova-Mirik, S., Elliott, N., Catana, V., Jones, D., Bowling, R., April 2006. Using digital image analysis and spectral reflectance data to quantify damage by green bug (Hemitera: Aphididae) in winter wheat. Computers and Electronics in Agriculture 51, 86–98.

Moya, E. A., Barrales, L. R., Apablaza, G. E., 2005. Assessment of the disease severity of squash powdery mildew through visual analysis, digital image analysis and validation of these methodologies. Crop Protection 24 (9), 785–789.

Nagel, H.-H., Summer 2004. Steps toward a Cognitive Vision System. AI Magazine 25 (2), 31–50.

Noordam, J., Hemming, J., van Heerde, C., Golbach, F., van Soest, R., Wekking, E., 2005. Automated Rose Cutting in Greenhouses with 3D Vision and Robotics: Analysis of 3D Vision Techniques for Stem Detection. Acta Horticulturae (ISHS) 691, 885–892.

Perner, P., Jänichen, S., Perner, H., 2006. Case-based Object Recognition for Airborne Fungi Recognition. Artificial Intelligence in Medicine 36 (2), 137–157.

Raviv, M., Blom, T., May 2001. The Effect of Water Availability and Quality on Photosynthesis and Productivity of Soilless-grown Cut Roses. Scientia Horticulturae 88 (4), 257–276.

Ross, D. A., Lim, J., Lin, R. -S., Yang, M. -H., 2007. Incremental Learning for Robust Visual Tracking. International Journal of Computer Vision, Special Issue: Learning for Vision (Online).

Skaloudova, B., Krivan, V., Zemek, R., September 2006. Computer-assisted Estimation of Leaf Damage caused by Spider Mites. Computers and Electronics in Agriculture 53 (2), 81–91.

Stern, V., Smith, R., Van den Bosch, R., Hagen, K., 1959. The integration of chemical and biological control of the spotted alfalfa aphid II. The integrated control concept. Hilgardia 29, 81–101.

Thonnat, M., Moisan, S., October 2000. What can Program Supervision Do for Software Reuse? IEE Proceedings - Software. Special Issue on Knowledge Modelling for Software Components Reuse 147 (5), 179–185.

Van Lanteren, J., Woets, J., 1988. Biological and integrated pest control in greenhouses. Annual Review of Entomology 33, 239–269.

Vernon, D., June 2006. The Space of Cognitive Vision. LNCS Vol. 3948 Cognitive Vision Systems, Part I: Foundations of Cognitive Vision Systems, pp. 7–24.

Vincze, M., Ponweiser, W., Zillich, M., 2006. Contextual Coordination in a Cognitive Vision System for Symbolic Activity Interpretation. In: ICVS'06 Fourth IEEE International Conference on Computer Vision Systems. Washington, DC, USA, pp. 12.

Xu, H., Ying, Y., Fu, X., Zhu, S., April 2007. Near-infrared Spectroscopy in detecting Leaf Miner Damage on Tomato Leaf. Biosystems Engineering 96 (4), 447–454.

Yang, Z., Rao, M., Elliott, N., Kindler, S., Popham, T., May 2005. Using ground-based multispectral radiometry to detect stress in wheat caused by greenbug (homoptera: Aphididae) infestation. Computers and Electronics in Agriculture 47 (2), 121–135.

Wang, N. and Zhang, N. and Wang, M., January 2006. Wireless sensors in agriculture and food industry–Recent development and future perspective. Computers and Electronics in Agriculture 50 (1), 1–14