Available at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper: AE—Automation and Emerging Technologies

Local feature-based identification and classification for orchard insects

Chenglu Wen[a,b], *Daniel E. Guyer*[b,*], *Wei Li*[a]

[a]*College of Engineering, China Agricultural University, Beijing 100083, China*
[b]*Department of Biosystems and Agricultural Engineering, Michigan State University, East Lansing, MI 48824, USA*

ARTICLE INFO

Article history:
Received 20 March 2009
Received in revised form
21 June 2009
Accepted 10 July 2009
Published online 28 August 2009

Insect monitoring in orchards under integrated pest management mainly relies on traps and manual evaluation. Automation, including machine vision combined with pattern recognition has achieved some applications in areas such as fruit sorting, robotic harvesting and quality detection, etc. An invariant local feature-based insect classification method has been proposed to automatically classify certain common insects in orchards. An invariant region feature detector was used to extract local features and a scale invariant feature transform (SIFT) descriptor was adapted to represent features obtained from the detector. To represent the whole image, the bag of words method was introduced to cluster and form a visual word expression for each of the image objects, which is normalized to feature vectors as input of the classifiers. Samples of five common pest species in orchards, *Cydia pomonella*, *Choristoneura rosaceana*, *Platynota idaeusalis*, *Argyrotaenia velutinana*, *Grapholita prunivora* were used to verify the classification method. Performances of six classifiers, which were minimum least square linear classifier (MLSLC), K nearest neighbour classifier (KNNC), Parzen density based linear classifier (PDLC), principal component analysis expansion linear classifier (PCALC), nearest mean classifier (NMC), and support vector machine (SVM) were compared by classification result. The best classification results under 10-fold cross-validation test were 4.57% and 5.95% using PCALC and SVM, indicating that the local region detector based insect classification method could be an effective way for insect identification and classification.

1. Introduction

Traditionally pest management in orchards has been accomplished by means of a regular spray programme which is based on a schedule rather than on the presence or likelihood of presence of insects in the field. More recently growers have incorporated weather-based models to predict pest and disease presence and apply control methods based on these models. The most accurate method to control pests, and a method which is gaining interest in

Abbreviations: ABIS, Automated Bee Identification System; DAISY, Digital Automated Identification System; GLOH, Gradient Location-Orientation Histogram; HSB, Hue Saturation Brightness; IPM, Integrated Pest Management; KNNC, K Nearest Neighbour Classifier; MLSLC, Minimum Least Square Linear Classifier; MSER, Maximally Stable Extremal Regions; NMC, Nearest Mean Classifier; PCA-SIFT, Principal Component Analysis- Scale Invariant Feature Transform; PCALC, Principal Component Analysis expansion Linear Classifier; PCBR, Principal Curvature-Based Region; PDLC, Parzen Density based Linear Classifier; RGB, Red Green Blue; SIFT, Scale Invariant Feature Transform; SPIDA, Species Identified Automatically; SUSAN, Smallest Univalue Segment Assimilating Nucleus; SVM, Support Vector Machine.

* *Corresponding author.*
E-mail addresses: kakapolu@msu.edu (C. Wen), guyer@msu.edu (D.E. Guyer), liww@cau.edu.cn (W. Li).

1537-5110/$ – see front matter © 2009 IAgrE. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.biosystemseng.2009.07.002

Nomenclature

$H(X)$	Hessian matrix of in the neighbourhood of one interest point
$I(X)$	Original image of an insect
$L_{xx}(X)$	The second partial derivative in x direction
$L_{xy}(X)$	The mixed partial second derivative in x and y direction
$L_{yy}(X)$	The second partial derivative in y direction
$L(X)$	Image smoothed by a Gaussian kernel $g(x, y, \sigma_1)$ for computing derivatives
σ_1	Scale value for smoothing Gaussian kernel
$g(x, y, \sigma_1)$	Gaussian kernel with scale σ_1
K	Numbers of clusters in K-means clustering method
N_m	Numbers of descriptor in m^{th} image in training set

the wake of the need to minimise environmental impacts, is integrated pest management (IPM). This is most commonly accomplished by monitoring using insect attractants and traps, such as light traps, bait traps and pheromone traps (Wise *et al.*, 2007). These traps must be monitored on a regular schedule and pest control operations are scheduled based on the findings in the traps.

Related work is presented in two parts; automated insect species identification and local feature-based object recognition.

1.1. *Automated insect species identification*

Growers monitor the environment in orchards by counting the harmful insects on the traps (usually sticky paper), and handle spraying according to the pest situation in orchards manually estimated by the insect distribution and population on the traps. The primary challenge with this process is the identification. Classification of insect species can be extremely time consuming and require technical expertise. Additionally, potential critical control timing could be missed due to time between trap checks. To minimise these problems, an automated insect recognition and classification method is needed.

Image-based insect recognition systems have been researched in recent years for various applications. Digital Automated Identification System (DAISY) can identify organisms within *Speciose arthropod* genera using a combination of both morphology and molecular data by matching principal component eigenimages of a training class (Weeks *et al.*, 1999). A Lucas continuous n-tuple method (Lucas, 1997) was used as a classifier in this system. DAISY is a genetic pattern matching system, which has additionally been applied to various types of image classification (Pajak, 2000; Watson *et al.*, 2003). Species Identified Automatically (SPIDA) (Russell *et al.*, 2005) is trained to identify the 121 species in the Australasian spider family *Trochanteriidae* based on an artificial neural network model. Automated Bee Identification System (ABIS) (Arbuckle *et al.*, 2001) can recognize bee species using geometric features of the forewing cells (such as, lengths, angles and areas) using a support vector machine(SVM) model and kernel discriminate analysis. Some research has moved forward to the recognition of the live insects. Mayo and Watson (2007) described different classifiers and datasets to identify live moths automatically, and indicated the subsets of the features matrix derived from binary and RGB/HSB (Red, Green, Blue/Hue, Saturation, Brightness) images are a way to reduce the size of a feature dataset and still get relatively high classification accuracy.

To extract features from the images, it is of great importance to get images of the insect where they adopt suitable poses. The need exists to obtain different views of an insect either by changing its pose manually or automatically using equipment. Larios *et al.*, (2008) developed a microscopy-based semi-automated mechanical manipulation and imaging system which captured the images of four stonefly taxa with various poses and a principal curvature-based region (PCBR) method for stonefly discrimination. Moreover, there is a trend in automated insect classification system development for researchers to try to identify and classify live insects since this allows for more physical characteristics to be retrained and this can result in images with additional features. Considering the complex situation related to closely related species, some methods have been developed to highlight areas of morphological difference between species, or amplify particular morphology, such as local feature analysis methods (Gaston and O'Neill, 2004), which is discussed in the following section.

1.2. *Local feature-based object recognition*

Features can be global and local. Global features based object recognition covers the entire image or a big portion of it, such as eigenspace matching (Hiroshi and Nayar, 1995), colour histograms (Swain and Ballard, 1991), and receptive field histograms (Bernt and Crowley, 1996). Because of their more global nature, they are not robust to occlusion and can suffer from a lack of invariance to similarity transformations such as scale or rotation, making them difficult to apply in practise. Local features are more robust to occlusion, background clutter, deformation and viewpoint change. Local features include edge, corner, entropy, curvature, region, ridge etc. Early work about local features can be traced back to Moravec's

Fig. 1 – Example images of (a) *Cydia pomonella*; (b) *Grapholita Prunivora*; (c) *Argyrotaenia velutinana*; (d) *Choristoneura rosaceana*; (e) *Platynota idaeusalis*.

Table 1 – Sample numbers and image numbers for different insects

	Cydia pomonella	*Choristoneura rosaceana*	*Platynota idaeusalis*	*Argyrotaenia velutinana*	*Grapholita Prunivora*	Total
Sample numbers	17	22	21	32	63	155
Image numbers	46	54	68	96	200	464

detector (1980). The algorithm firstly detects the corner, and then uses the sum of squared differences to measure similarity between two nearby patches. Region detectors can find the interest region and points in the image, and includes Harris corner detector (Harris and Stephens, 1988), Harris-affine, Hessian-affine detector (Mikolajvzyk and Schmid, 2004), maximal stable extremal regions (Matas *et al.*, 2002), Kadir detectors (Kadir *et al.*, 2004), and edge based region detector and intensity extrema detector (Tuytelaars and Van Gool, 2004). The six latter methods are affine invariant region detectors which mean the local features detected by these detectors can be not only invariant to translations, rotations, uniform rescaling in the spatial domain, but also affine transformations.

Local feature representation for the detected region is the follow-on step of object recognition algorithm. Recently, some new feature representation methods have been developed. Lowe (1999, 2004) developed a 128 dimensional scale invariant feature transform (SIFT) descriptor. A gradient location-orientation histogram (GLOH) (Mikolajczyk and Schmid, 2005) is based on SIFT descriptor, but it computes the SIFT descriptor for a log-polar location grid with 17 location bins. Shape context is a three-dimensional (3-D) histogram of edge point locations and orientations, with the edge information extracted by the Canny edge detector. The Principal Component Analysis (PCA)–SIFT (Ke and Sukthankar, 2004) descriptor is a 36 dimensional vector which is reduced from a 39 × 39 location gradient region computed by SIFT with the PCA technique. Spin image representation (Lazebnik *et al.*, 2003) is a histogram of quantized pixel locations and intensity values, where the intensity of

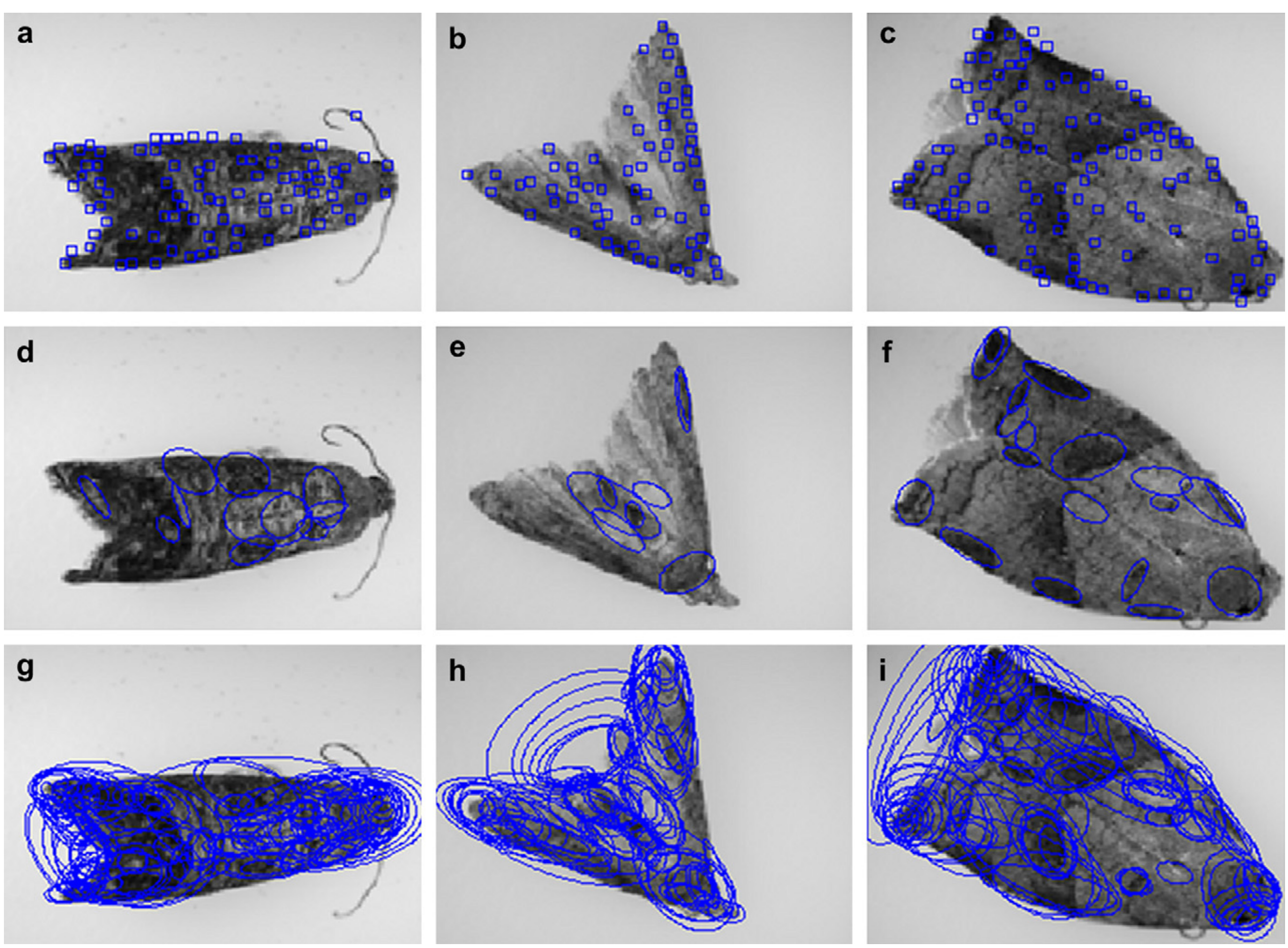

Fig. 2 – Interest points and regions detected by (a)–(c) SUSAN corner detector; (d)–(f) MSER detector; (g)–(i) Hessian-affine detector.

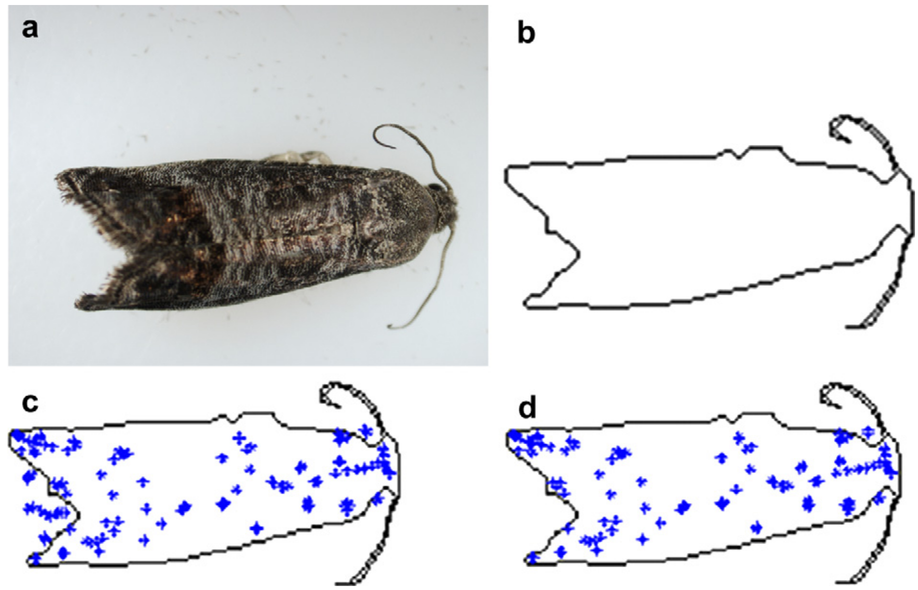

Fig. 3 – Interest points refined by morphological segmentation result. (a) Original image of C. *Pomonella*; (b) Contour of the insect after morphological segmentation; (c) Interest points overlap on (b); (d) Resulting interest points after eliminating the points out of the boundary of insect.

a normalized patch is quantized into 10 bins. Other descriptors also have been used to describe the feature region, like steerable filters (Freeman and Adelson, 1991), moment invariants (Van Gool *et al.*, 1996), cross correlation, etc. For certain experimental situations, SIFT and SIFT-based descriptors give the highest matching accuracy with scale, translation, rotation and affine change, and best performance on distinctiveness and robustness of descriptors (Mikolajczyk and Schmid, 2005).

The focus of our initial work was automated classification of harmful insects in cherry and apple orchards, which currently are manually classified by experts with entomological knowledge. Automated classification of the insects to harmful or harmless types, and population counting of each type, can establish baseline criteria for evaluation of the ecology environment in the orchard and monitoring the production of the fruit. The paper aims to develop a local feature-based method to identify common insects in cherry and apple orchards, and address the problem of automatically classifying the insects.

It should be noted that to a trained eye of an expert, insect identification is not challenging, especially when a finite insect domain is involved. The overall and long-term goal of the project is to establish continuous or semi-continuous monitoring and reducing the requirement for expert labour. Thus, the target of this effort is not to have improved insect recognition over current manual methods but rather to take a step towards automation with equivalent accuracy. An intermediate step towards full automation is to develop a system capable of 'reading' or evaluating the insects on traps currently in practise with the thought automation of continuous monitoring or automated collection or checking of traps

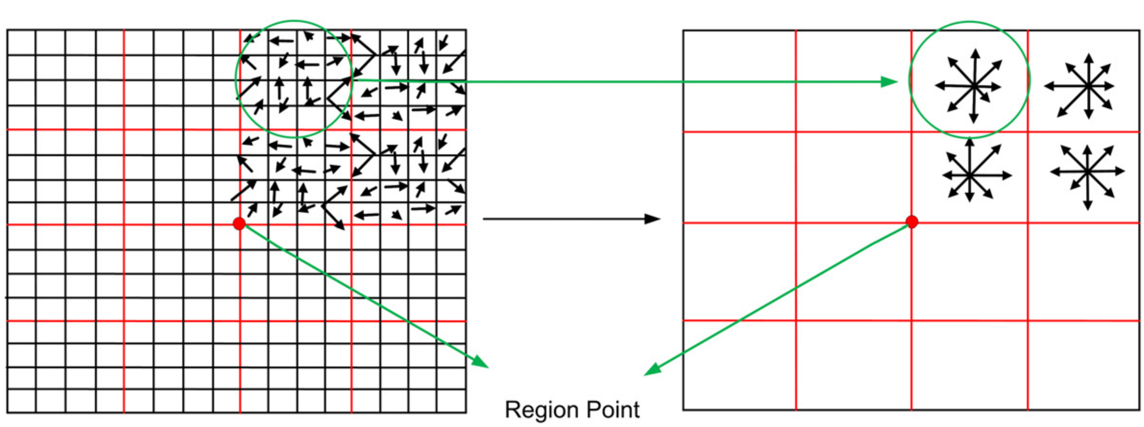

Fig. 4 – A 4 × 4 SIFT descriptor array is computed from a 16 × 16 set of pixel points, which has gradient magnitude and orientation (Lowe, 1999).

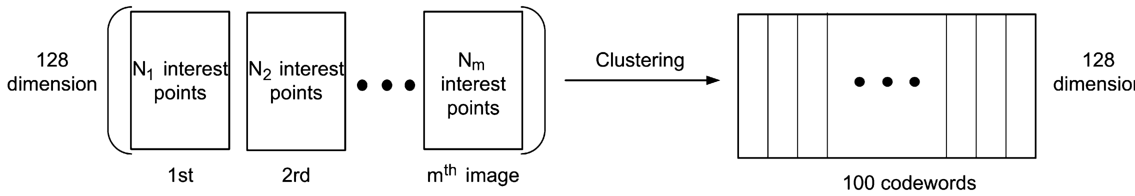

Fig. 5 – Formation of a codeword dictionary from all interest descriptors.

would follow in future development. An image-based insect identification system is a potential tool in both the intermediate trapping system and in the broader automation goal.

2. Materials and method

2.1. *Image acquisition*

Five insect species, *Cydia pomonella* (codling moth), *Grapholita Prunivora* (lesser appleworm), *Argyrotaenia velutinana* (redbanded leafroller), *Choristoneura rosaceana* (obliquebanded leafroller), and *Platynota idaeusalis* (tufted apple budmoth), were obtained from the IPM pest colonies at Michigan State University, USA. The insects were frozen for 20 min and then randomly placed on a white balance panel under the reflectance light base of a Nikon stereoscopic zoom microscope SMZ1000 (Nikon, Tokyo, Japan) with Plan Apochromat 0.5 × objective. A DS-Fi1 colour digital camera (Nikon, Tokyo, Japan) was mounted on the microscope. Illumination is provided by a gooseneck light guide powered by a Schott-Fostec Eke Pheostat 150 W light source (Schott North America Inc., NY, USA). The images were taken under different insect orientations and two poses (top and side view), which were found to be the most common poses in sticky traps collected from field for IPM. Some examples of insect images are shown in Fig. 1, and the numbers of samples and images used in experiments are given in Table 1.

2.2. *Insect local feature detectors*

A feature detector is commonly used to extract stable and informative regions from images in order to reduce the computational complexity and improve the robustness to image deformation.

As a semi-segmentation method, local feature detectors have become increasingly popular over the last few years. Local feature-based detectors overcome the need for semantic-level segmentation. Separating the different foreground objects from the background is a difficult problem. Representing the image as a set of overlapping local regions yields an implicit segmentation. Since the features are local, some of them cover part of the foreground object and can be considered as relevant, while others fall on the background or on object boundaries and can be considered as irrelevant.

The corner detector, originally referred to as a Harris corner detector, detects the local points with two directional

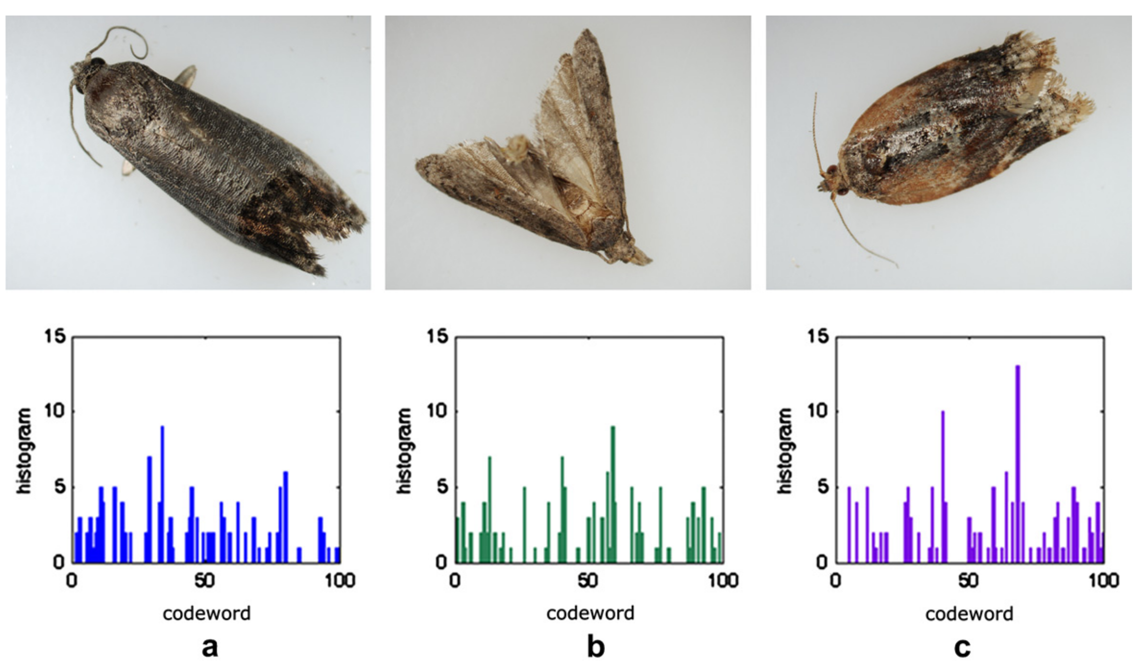

Fig. 6 – Feature histogram representation for (a) *Cydia pomonella*; (b) *Platynota idaeusalis*; (c) *Argyrotaenia velutinana* using bag of words representation method.

Table 2 – Classification accuracy from cross validation with 10 times repetitions (%)

	MLSLC	PDLC	KNNC	PCALC	NMC	SVM
Average test error	7.2	6.77	7.67	4.57	9.4	5.95
Standard deviation over the repetitions	0.35	0.67	0.51	0.41	0.32	0.37

information; differing from edge detection. The corner detector has been replaced by an interest point detector now which not only shows the edge intersection but the point brighter or darker than its neighbours. An example is the smallest univalue segment assimilating nucleus (SUSAN) point detector (Smith and Brady, 1997).

Unlike corner detection, regions of interest provide a description of image structures in terms of regions, not points, and can detect areas which are too smooth to be detected by corner detectors. From Mikolajczyk's work (2005), performances of some region detectors have been evaluated against changes in viewpoint, scale, illumination, defocus and image compression. For most experiments the maximally stable extremal regions (MSER) and the Hessian-affine detector obtained the best repeatability score, so the detection results of these two detectors were compared on insect samples. The detection results of SUSAN, MSER and Hessian-affine detectors for different species are shown in Fig. 2.

A MSER detects regions through thresholding instead of filtering. The pixels inside the MSER region have either higher or lower intensity than the pixels on its outer boundary. The MSER region is stable over a large range of thresholds. The maximal stability is measured by the relative area change as a function of thresholds.

The Hessian-affine detector uses a multiple scale iterative algorithm to spatially localize and select the affine invariant points and scale. At each individual scale, the Hessian-affine detector chooses the interest points based on the local maxima of the Hessian matrix in the neighbourhood of that point:

$$H(X) = \begin{bmatrix} L_{xx}(X) & L_{xy}(X) \\ L_{xy}(X) & L_{yy}(X) \end{bmatrix} \quad (1)$$

where $L_{xx}(X)$ is the second partial derivative in x direction, $L_{yy}(X)$ is the second partial derivative in y direction, and $L_{xy}(X)$ is the mixed partial second derivative in two directions. The derivatives are computed in the current iteration scale and are derivatives of an image smoothed by a Gaussian kernel: $L(X) = g(x, y, \sigma_1) \otimes I(X)$, where $I(X)$ is the original image, and $g(x, y, \sigma_1)$ is the Gaussian Kernel:

$$g(x, y, \sigma_I) = \frac{1}{2\pi\sigma_I^2} e^{-(x^2+y^2)/2\sigma_I^2} \quad (2)$$

The extrema of the Laplacian of Gaussian of the image over scalar space is used to select scale (Lindeberg, 1998). The affine elliptical region is iteratively estimated by the second moment matrix of the intensity gradient which gives gradient distribution in a local neighbourhood of a point (Lindeberg and Garding, 1997).

The Hessian matrix-based detector gives strong responses on regions because of the second derivatives in the matrix. From Fig. 2, the Hessian-affine detector provides better description due to its more complete covering of the whole insect and more local information extracted than MSER and SUSAN detectors, and thus the result of Hessian-affine region detector will be used as input of our feature descriptor.

The interest region detected is around one interest point. Because the interest points are the result of responses to local maxima of intensity or gradient, some of these points are located on the background. Here the morphological segmentation result was introduced to eliminate the low contrast points, and only the intensity points inside the insect boundary are kept (As shown in Fig. 3). This step was only required if it is desired to remove any background features identified in the local feature extraction process.

2.3. *Feature descriptors*

A feature descriptor is used to describe the local region of interest. The simplest feature descriptor is a vector of all pixels within an image, however, the high dimensionality of such a description results in high computational complexity for recognition. In recent years, some new techniques have been developed to reduce the high dimension.

The SIFT obtains the best results in most applications. SIFT is a 3-D histogram of gradient location and orientation, where a 128-element SIFT vector is used to describe the local neighbourhood of each point of interest, representing the 4×4 location grid of 8-bin orientation histograms for each interest point detected from local region detector. The scale of each vector is the weighted value of the histogram of the pixel's gradient vector (shown in Fig. 4). The SIFT descriptors were applied to describe the regions of interest.

2.4. *Bag of words image representation*

For the descriptors obtained above, the normalisation required them to be converted to generalised vectors to represent the image. The bag of visual words (features) method is introduced to convert each set of vectors (with label) to a standard feature vector by adding these same label feature points into a histogram (Li and Perona, 2005). The K-means clustering method was used to cluster all the descriptors of the training set into K object parts (codewords) to form a codeword dictionary. The result of K

Table 3 – Image numbers used for training and testing

Species name	*Cydia pomonella*	*Choristoneura rosaceana*	*Platynota idaeusalis*	*Argyrotaenia velutinana*	*Grapholita Prunivora*	Total
Training set	27	31	40	60	116	274
Testing set	19	23	28	36	84	190

Table 4 – Classification accuracy of six classifiers

Species	Sample number	MLSLC	PDLC	KNNC	PCALC	NMC	SVM	Average rate, %
Cydia pomonella	19	11	14	12	11	17	14	69.3
Choristoneura rosaceana	23	17	18	12	20	18	19	75.4
Platynota idaeusalis	28	23	24	19	23	23	25	81.5
Argyrotaenia velutinana	36	25	26	20	28	31	28	73.1
Grapholita Prunivora	84	84	82	84	83	81	82	98.4
Total number	190	160	164	147	165	170	168	85.4
Correct rate, %		84.2	86.3	77.4	86.8	89.5	88.4	

with value of 50 and 100 were compared, and the value of 100 was chosen for its better performance. For each image, output of the feature descriptor was a matrix with the size of numbers of interest points by 128-element feature vector. The feature descriptor for the training set was a $(N_1 + N_2 + \ldots N_m)$ by 128-element matrix, which was clustered into a 100 by 128 matrix codeword dictionary (As shown in Fig. 5).

For a training image with N_p interest points, the distance is calculated between the 128-element features of each interest point and the codeword dictionary, and assigns each interest points to a certain codeword group with the smallest distance. For a training image, the number of the point of interest with the same codeword are counted and distributed into a 100-bin histogram, which were normalised features extracted for classification. Examples of feature histogram results for different species are shown in Fig. 6.

3. Results and discussion

Six classification methods which included minimum least square linear classifier (MLSLC), Parzen density based linear classifier (PDLC), K nearest neighbour classifier (KNNC), principal component analysis based linear classifier (PCALC), nearest mean classifier (NMC), and SVM with polynomial kernel, were implemented and compared (Duin *et al.*, 2004). Here, the KNNC and NMC, which are distance-based classifier, used χ^2 (Chi-squared) distance for histogram distance measurement (Rubner *et al.*, 2001).

3.1. Overall result by cross-validation

For each feature in an image was normalized to a 100-bins histogram, a cross-validation method is more convincible to evaluate the classifier. A 10-fold cross-validation with 10 times repetitions was used to test the sample and compare the performances of these six classifiers. The results of the cross-validation experiment are shown in Table 2. The results show the PCALC and SVM methods outperformed than other classifiers, and the distance-based classifiers yielded poor performance.

3.2. Insect misclassification analysis

To analyse the insect misclassification situations, the 464 images from five species were randomly divided into a training set and a testing set. The number of the training images was 274 and the testing image set was 190 (Number of each species are shown in Table 3).

To obtain the performance of each classifier they were run with the testing set, and the results of classification are shown in Table 4.

For all species combined under individual classifiers, we found the NMC and SVM to show the best performance on overall classification, with 89.5% and 88.4% of the testing samples correctly classified. Following this were PCALC and PDLC, with correct classification rates of 86.8% and 86.3%. MLSLC correctly classified 84.2% of the tested samples, and the poorest classifier was KNNC, which yielded a correct classification rate of 77.4%.

Looking at individual species across different classifiers, it is clear that *G. Prunivora* has the highest correct classification rate, meaning it is relatively easily recognized. The *C. pomonella* has a relatively low correct classification rate. The possible reason for this is the number of training images for *G. Prunivora* is nearly four times number of the *C. pomonella*. It trend is that larger training datasets can improve the accuracy of classification.

To analyse the misclassification results of different species, SVM which gave relatively high performance in cross-validation and testing set classification experiment, was selected and analysed. The confusion matrix of SVM is shown in Table 5.

Table 5 – Confusion matrix of the SVM result

Predicted as Test species	*Cydia pomonella*	*Choristoneura rosaceana*	*Platynota idaeusalis*	*Argyrotaenia velutinana*	*Grapholita Prunivora*
Cydia pomonella	14	1	1	3	0
Choristoneura rosaceana	0	19	0	4	0
Platynota idaeusalis	2	0	25	0	1
Argyrotaenia velutinana	1	0	3	28	4
Grapholita Prunivora	0	0	2	0	82

From Table 5 it was found that in the SVM classification, several samples of the three other species except *G. Prunivora* were misclassified to *A. velutinana*. Tracing back to original testing images to find potential answers, it was discovered that two of *C. pomonella* were misclassified as *A. velutinana* because the images of these two specific samples had more reflected light from their wings than most others, which made some of their local textures features invisible. A possible reason for four *C. rosaceana* being misclassified as *A. velutinana* was the wings of these insects opened wider than normal samples, which decreased the features from their wings. Some *P. idaeusalis* samples were misclassified as *C. Pomonella* since the part of the features were lost because of the side view of the samples.

In our research, the useful features for identification and classification are mainly on the wings of insects. However, the views of wings varied with body pose (i.e., top or side view) and wing pose (i.e., open or closed status), which brings more difficulties and classification errors.

An issue of possible concern was how to identify insects when they could potentially present themselves for viewing in a variety of pose and angle combinations. However, whilst working with in-field trapping systems, it was noted that, at least initially, insects tend to land in the traps in a resting-like position with feet down and wings closed. This greatly reduces the potential range of poses that the insects might present to the imaging system.

Most past and current insect identification studies appear to have focused on insects of very close taxonomy for the task they presented. In the automated pest detection task for orchard monitoring the challenge exists to differentiate several insect species, possibly in the same family and under various poses, such as with wings closed or open, or laying on one side. However, the likelihood of attracting a certain species is due to the pheromone attractant used. These identification challenges were addressed in this study by including realistic samples with varying poses and moths from a similar family. It is evident that classifying more dissimilar insects, such as moths *versus* flies or beetles would be a less complicated system.

4. Conclusions

This paper summarised insect species classification using local region feature detectors for feature detection, the SIFT region descriptor for local feature description, bag of visual words method for image representation, and six classifiers for classification. The work aimed to determine whether local feature-based classification could be effective for orchard insect species classification. Because the insect species used in our research were all moths, the challenge is similarity of inter-species (e.g. when insect wings are closed this makes many of their unique wing patterns invisible), and the intra-species variation (e.g. the size difference between the female and male *C. pomonella*).

Local features overcome some disadvantages of global features, such as partial or total invariance to rotation, illumination, scale, and occlusion and/or segmentation dependence. Local features based insect species classification also provides a good example of non-segmentation classification. The classification results indicated that local features are effective for insect classification. The PCALC and SVM methods achieved the highest accuracy on five insect species by 10-fold cross-validation test. A disadvantage of this local feature-based insect classification method is that global information related to the whole object is not considered and these are very important factors for entomologists who wish to identify and classify the insects. Future research for this work should combine global features and local features to obtain higher classification accuracy.

Acknowledgements

The authors would like to thank the support from USDA-ARS Postharvest Engineering Lab at East Lansing, Michigan. The work was partly supported by the China Scholarship Council, Chinese Government Scholarship Program.

REFERENCES

Arbuckle B; Schroeder S; Steinhage V; Wittmann D (2001). Biodiversity informatics in action: identification and monitoring of bee species using ABIS. In: Proceedings of 15th International Symposium for Environmental Protection, Zurich, pp 425–430.

Bernt S; Crowley J L (1996). Object recognition using multidimensional receptive field histograms. In: Fourth European Conference on Computer Vision, Cambridge, UK, pp 610–619.

Duin R P W; Juszczak P; Paclik P; Pekalska E; De Ridder D; Tax D M J (2004). PRTools4: A matlab toolbox for pattern recognition, Delft University of Technology.

Freeman W; Adelson E (1991). The design and use of steerable filters. IEEE Transaction on Pattern Analysis and Machine Intelligence, **13**(9), 891–906.

Gaston K J; O'Neill M A (2004). Automated species identification: why not? Philosophical Transactions of Royal Society B: Biological Sciences, **359**(1444), 655–667.

Harris C; Stephens M (1988). A combined corner and edge detector. In: Proceedings of the Fourth Alvey Vision Conference, pp 147–151.

Hiroshi M; Nayar S K (1995). Visual learning and recognition of 3-D objects from appearance. International Journal of Computer Vision, **14**(1), 5–24.

Ke Y; Sukthankar R (2004). PCA–SIFT: a more distinctive representation for local image descriptors. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, pp 506–513.

Kadir T; Zisserman A; Brady M (2004). An affine invariant salient region detector. In: European Conference of Computer Vision, pp 228–241.

Larios N; Deng H; Zhang W; Sarpola M; Yuen J; Paasch R; Moldenke A; Lytle D A; Correa S R; Mortensen E; Shapiro L G; Dietterich T G (2008). Automated insect identification through concatenated histograms of local appearance features. Machine Vision and Applications, **19**, 105–123.

Lazebnik S; Schmid C; Ponce J (2003). A sparse texture representation using affine-invariant neighborhoods. In: Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, **2**, pp 319–324.

Li F; Perona P (2005). A Bayesian hierarchical model for learning natural scene categories. In: Proceedings of the 2005 IEEE

Computer Society Conference on Computer Vision and Pattern Recognition, **2**, pp 524–531.
Lindeberg T; Garding J (1997). Shape-adapted smoothing in estimation of 3-d shape cues from affine deformations of local 2-d brightness structure. Image and Vision Computing, **15**(6), 415–434.
Lindeberg T (1998). Feature detection with automatic scale selection. International Journal of Computer Vision, **30**(2), 79–116.
Lowe D G (1999).Object recognition from local scale-invariant features. In: Proceedings of International Conference on Computer Vision, pp 1150–1157.
Lowe D G (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, **60**(2), 91–110.
Lucas S (1997). Face recognition with continuous n-tuple classifier. In: Proceedings of British Machine Vision Conference, pp 222–231.
Matas J; Chum O; Urba M; Pajdla T (2002). Robust wide-baseline stereo from maximally stable external regions. In: Proceedings of British Machine Vision Conference, pp 384–393.
Mayo M; Watson A T (2007). Automatic species identification of live moths. Knowledge-Based Systems, **20**, 195–202.
Mikolajvzyk K; Schmid C (2004). Scale and affine invariant interest point detectors. International Journal of Computer Vision, **60**(1), 63–83.
Mikolajczyk K; Schmid C (2005). A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, **27**(10), 1615–1630.
Moravec H (1980). Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover. Tech Report CMU-RI-TR-3 Carnegie-Mellon University, Robotics Institute.
Pajak M (2000). Identification of British Bombus and Megabombus using DAISY, B.A. 3rd Year Honours Project, Oxford.
Rubner Y; Puzicha J; Tomasi C; Buhmann J M (2001). Empirical evaluation of dissmilarity measures for color and texture. Computer Vision and Image Understanding, **84**, 25–43.
Russell K N; Do M T; Platnick N I (2005). Introducing SPIDA-web: an automated identification system for biological species. Taxonomic Database Working Group Annual Meeting.
Smith S M; Brady J M (1997). SUSAN – a new approach to low level image processing. International Journal of Computer Vision, **23**, 45–78.
Swain M; Ballard D (1991). Color indexing. International Journal of Computer Vision, **7**(1), 11–32.
Tuytelaars T; Van Gool L (2004). Widely separated views based on affine invariant region. International Journal of Computer Vision, **59**(1), 61–85.
Van Gool L; Moons T; Ungureanu D (1996). Affine/Photometric invariants for planar intensity patterns. In: Proceedings of Fourth European Conference on Computer Vision, pp 642–651.
Watson A T; O'Neill M A; Kitching I J (2003). Automated identification of live moths (*Macrolepidoptera*) using igital automated identification system (DAISY). Systematic and Biodiversity, **1**, 287–300.
Weeks P J D; O'Neill M A; Gaston K J; Gauld I D (1999). Automating insect identification: exploring the limitations of a prototype system. Journal of Applied Entomology, **123**, 1–8.
Wise J C; Gut L J; Isaacs R (2007). Michigan Fruit Management Guide. Department of Plant Pathology and Department of Horticulture, Michigan State University.