HGCA

PROJECT REPORT No. 156

RAPID AUTOMATED DETECTION OF INSECTS AND CERTAIN OTHER CONTAMINANTS IN CEREALS

MARCH 1998

Price £5.00

PROJECT REPORT No. 156

RAPID AUTOMATED DETECTION OF INSECTS AND CERTAIN OTHER CONTAMINANTS IN CEREALS

by

J CHAMBERS[1], C RIDGEWAY[1], E R DAVIES[2], D R MASON[2] and M W BATEMAN[2]

[1] Central Science Laboratory, Sand Hutton, York YO4 1LZ

[2] Machine Vision Group, Department of Physics, Royal Holloway, University of London, Egham, Surrey TW20 0EX

This is the final report of a twenty-seven month project which started in March 1995. The work was funded by a grant of £121,854 from the Home-Grown Cereals Authority (project No 0048/1/94).

RAPID AUTOMATED DETECTION OF INSECTS AND CERTAIN OTHER CONTAMINANTS IN CEREALS

Reference: 0048/1/94

J. Chambers[1]*, C. Ridgway[1], E. R. Davies[2], D. R. Mason[2] and M. W. Bateman[2]

[1]Central Science Laboratory, Sand Hutton, York, YO4 1LZ
[2]Machine Vision Group, Department of Physics, Royal Holloway, University of London, Egham, Surrey, TW20 0EX
*to whom correspondence should be addressed.

Final Report

Duration: 27 months (including 3 month hiatus) Starting Date: 1 March 1995

Report completed: 29 August 1997

ABSTRACT

The research reported here provides a basis on which to develop the first automated system for detecting contaminants in post-harvest grain which will meet trade requirements in terms of price, reliability and speed. This has been done by developing analytical methods for images of wheat samples containing various contaminants.

A suitable system to capture images has been developed. The most promising methods for analysis of the images were found to be thresholding and linear feature detection but each needed substantial innovative modification to deliver the required performance. Application of these approaches to images recorded in the visible region has shown that it will be possible to detect and distinguish contaminants into the following categories: (a) adult and larval insects external to grain kernels, (b) ergot and droppings from rats and mice, (c) rapeseeds, (d) large dark objects, and (e) other contaminants. The system would not report as contaminants permitted admixture such as hulls and damaged grains. Reliability of detection and recognition varied with type of contaminant from complete success with sets of images containing a total of 60 items of either ergot or rapeseed to 6 missed items out of 150 larval *O. surinamensis.* To minimise false alarms, it is recommended that each image does not exceed 25 grains. Painstaking improvements to the software have brought a substantial reduction in time needed to analyse each image and the application of faster computers soon to be available should allow the system to handle samples at the required rate of 3 kg in 3 minutes. The cost of the computing hardware, excluding that required for sample handling, is likely to be around £4000.

No existing device has a performance comparable to that expected from the system to be developed from the work presented here and a suggested specification for it has been written. The performance figures it contains should be regarded as starting points for further improvement, not the best ultimately achievable. Attention is now being given to protection of the intellectual property upon which this system depends and the identification of a suitable vendor to bring the system to market.

Internally infested grains cannot be detected by imaging in the visible region. Incorporation of an X-ray camera into the system was investigated but this would increase cost, complexity, and time needed, so its inclusion is not justified. However, such infested kernels were detectable from images recorded at carefully chosen wavelengths in the NIR region and the incorporation of an NIR capability into the inspection system should permit the reliable and rapid detection of internal infestation.

CONTENTS

OBJECTIVE

The main objective of this project was to conduct research on which to base development of the first automated system for detection of contaminants in post-harvest grain which would meet trade requirements in terms of price, reliability and speed. Building on the results of previous research, the intention was to investigate the potential of image analysis by automated visual inspection. This was to be applied to the detection of non-insect contamination and insect contamination, followed by integration of these capabilities into a single system. The intended output was planned to be a specification by which proof of theories in the laboratory could be translated into a prototype machine for testing under practical conditions.

INTRODUCTION

Increasingly competitive markets, especially internationally, mean that the cereal trade must deliver products of ever higher quality. Contamination of harvested cereals by insects, mites and extraneous matter such as ergot continues to cause problems but there is still no satisfactory method for their rapid detection for grain in transit. With new technology there is a real chance of overcoming this difficulty.

Previous work funded by the HGCA investigated the potential of near infra-red spectroscopy (NIR) (Chambers *et al.*, 1994). This showed that NIR can detect insects of different species and of different developmental stages, in different varieties of wheat at different moisture contents. This work proved the feasibility of developing a rapid practical method which could be automated but indicated that NIR could only differentiate between uninfested samples and those with about 0.27 insects g^{-1} or more. The purpose of the current project was to apply new developments which would allow this limit to be reduced substantially so that no contamination would be missed.

The new developments are based on the recording and analysis of images of the grain. The Machine Vision Group at Royal Holloway, University of London (RHUL) has long been expert in vision algorithms and their application to food inspection (Davies, 1993; 1996). The group has investigated entirely new computer vision-based methods for detecting foreign bodies in foodstuffs. Their work has given impressive results in the detection of inanimate objects (e.g. metal, stone, wood, plastic and glass) in food packets of sweetcorn, peas or

chocolate (Patel *et al.*, 1993; Hannah *et al.*, 1995). This experience has been exploited in the present project for the development of new methods which can detect contamination in grain.

The application of machine vision methods in determining grain quality has been investigated by various workers in recent years. Efforts have mostly concentrated on the determination of varietal purity (Keefe, 1992), but detection of foreign material (Zayas and Steele, 1990) and damaged grains (Liao *et al.*, 1994) have also been demonstrated. In carrying out this work it has been anticipated that the continuing development of ever-higher performance computers and cameras would allow a useful (statistically significant) amount of grain to be analysed in a given time and at affordable cost. However, existing devices still fall short of the performance and affordability required by most UK cereal traders. For example, the semi-automatic wheat kernel shape analyser developed by Keefe *et al.* (1992) took 5 minutes to analyse 50 grains and the real-time maize kernel defect detector developed by Liao *et al.* (1994) analysed twelve kernels in around two seconds. A commercially-available device, the Tecator 310 GrainCheck, analyses less than 100 g grain in 2-3 minutes for varietal purity, damaged kernels, foreign seeds and other foreign material. The price is presently around £50,000. This type of device may be very useful in limited specific applications such as plant breeding or seed inspection but there is a need for a much faster and much less expensive device which analyses quality parameters of most importance to the average grain trader.

The consensus view of various members of the HGCA R&D Committee at a meeting with the project collaborators was that the device should have the following characteristics:

- be inexpensive; the cost would have to be as low as possible (a few thousand pounds) if it was to be bought by dealers and processors
- detect rodent droppings, ergot, adult and larval insects external to grain kernels
- detect hidden infestations if it could be achieved at low cost
- distinguish between different classes of contaminant and signal that a particular type of contaminant had been found
- be capable of preparing and screening samples of up to 3 kg automatically
- be rapid in its performance, although speed required would depend on needs of different end users.

General

Except where stated otherwise, samples were prepared, images recorded and the quality of the images assessed at CSL, while image analysis was undertaken at RHUL. Image analysis uses a number of specific terms. A glossary of the terms of principal importance is provided as Appendix A.

Imaging in the visible region

Materials used in the images

Control wheat used was variety Mercia at 14% moisture content or [Experiments (g) and (h) only] variety Hereward at 14% moisture content, cleaned by the supplier. Control rapeseed used was of variety Capricorn at around 8% moisture content. Control barley used was of variety Halcyon at 16% moisture content. The laboratory insects used were *Oryzaephilus surinamensis* (saw-toothed grain beetle), *Sitophilus granarius* (grain weevil), *Tribolium castaneum* (rust-red flour beetle), *Rhizopertha dominica* (lesser grain borer), *Ahasverus advena* (foreign grain beetle) and *Cryptolestes ferrugineus* (rust-red grain beetle), all of the Slough Laboratory insecticide susceptible strains. Adult insects were aged approximately 0-4 weeks after emergence. The laboratory larvae used were *O. surinamensis* as above, aged approximately 2-4 weeks after egg-laying. These adult and larval insects were freshly-killed by chilling in solid carbon dioxide for ten minutes. Rat droppings were obtained from a Slough Laboratory strain of *Rattus norvegicus* and left to dry out on trays in a fume-cupboard over two days before use. Dry mouse droppings were obtained from a Slough Laboratory strain of *Mus domesticus* and used as received. Ergot used in initial algorithm development was supplied as a demonstration sample from a commercial supplier of rice-sorting equipment and used as received [from rye; *ergot (1)*]. Two other samples of ergot, isolated from commercial grain samples [one from wheat, *ergot (2)*; one from rye, *ergot (3)*], were obtained *via* the HGCA. Two bags of UK wheat (approximately 500 g each; variety and moisture content unknown) containing permitted admixture (Table 1) were obtained from RHM Technology Limited. Two bags of wheat (approximately 1 kg each; variety and moisture content unknown) heavily infested with insects (Table 1) were obtained from Igrox Limited.

Sample presentation

Monolayer wheat samples in which the kernels were touching were prepared by adding around 100 g wheat to a tray of dimensions 30 cm x 14.5 cm, containing at its centre a square of white card of dimensions 10 cm x 10 cm x 1 mm thickness. The area of the card was slightly larger

than that of the acquired image (6.5 cm x 6.5 cm) and the card was fixed to the tray using tape. Careful sliding of the grain from side to side on the tray produced a monolayer of wheat over the entire area of the card square. Monolayer wheat samples in which the kernels were not touching were prepared in a similar way, but adding only around 30 g wheat to the tray and carefully agitating to achieve a reasonable separation between most kernels. In the case of artificial (laboratory-generated) samples, any foreign objects already present in the cleaned wheat sample were located by visual inspection and removed manually. Contaminant items were then dropped at random locations within the area of the wheat sample seen by the camera, followed by gentle manual agitation of the tray to settle the sample. Where the kernels were touching, the amount of wheat included in the image obtained by the camera system in a typical sample weighed 9.5 g. Where the kernels were not touching, the amount was estimated to be in the range 1 - 1.5 g.

Method of image capture

The general arrangement for the capture of images is shown in Figure 1. Monochrome digital images of resolution 256 x 256 pixels x 8 bit (256 grey levels) were obtained using an Ikegami ICD-42E Type F 1/2 inch CCD camera and Apple Macintosh Quadra 700 computer fitted with frame grabber and NIH Image v1.51 operating software. The camera was fitted with a Fujinon 17.5-105 mm TV zoom lens set at 30 mm focal length and positioned 107 cm above the sample. This combination of zoom setting and distance from sample was selected from several investigated (17.5-30 mm and 87-107 cm respectively) and enabled the maximum size of wheat sample to be scanned at acceptable resolution. Lighting was from a 40 W circular fluorescent tube of 34 cm internal diameter, held 34 cm above the sample. The fluorescent tube and controlling electronics were mounted to the underside of a wooden board of dimensions 52 cm x 45 cm with a central circular hole of diameter 29 cm. The TIFF image files obtained from this system were converted to MS-DOS format and stored on micro-floppy disk.

Images recorded: laboratory-generated samples

Sets of images recorded are listed in Table 2. Except where stated otherwise, one image was recorded for each of ten replicate samples at each of six levels of contamination, these containing 0, 1, 2, 3, 4, or 5 items of contaminant. Thus for each type of contaminant (a) to (f), a set of 60 images was recorded. For (g), one image was recorded for each of twenty replicate samples in which one item of contaminant was present, repeated for each of the three different ergot samples. Here, images were also recorded for 20 replicate samples containing no contaminant, giving a set of 80 images total. For (h), one image was recorded for each of

six replicate samples containing no contaminant. For each insect species one sample, in which five items of contaminant were present, was imaged, giving a set of 12 images total.

Duplicate sets of images (a), (c) and (d) were also saved, with the location of the contaminant marked by the person who had prepared the sample, using the NIH Image software. These were used to aid the operator in setting up the image analysis system.

Images recorded: authentic samples

For (j) and (k), one image was recorded for each of 10 samples extracted at random from the total sample of about 500 g, for each experiment. As only a small fraction of the approximately 30 g samples were imaged, any admixture particulates of interest falling outside the imaged area were moved to within this area. Admixture particulates present in each recorded image were identified by visual inspection and listed for reference. The same method was used for (l) and (m) except that here one image was recorded for each of 30 samples extracted at random, from the total sample of about 1 kg, for each experiment.

Quality of the recorded images

The 50 wheat images from experiment (a) above which contained at least one barley kernel were displayed on the computer to an individual not directly involved in the acquisition of the images. The individual was asked to identify and mark the positions of the barley kernels in each image, taking as much time as required. As an aid, the images could be contrast enhanced and sharpened as required using the NIH Image software (for example Figure 2). These inspected images with the suspected barley kernels marked were then compared to the duplicate set of images in which the actual positions had been marked by the person who had prepared the sample (Figure 3). The percentage of barley kernels successfully identified by the individual was then calculated.

Methods of image analysis

As an initial familiarisation experiment, the sixty images in Experiment (b), rapeseed in wheat, were analysed by a simple thresholding algorithm preceded by a process known as minimisation, which involves use of a filter to enhance contrast between the contaminant and the cereal background.

Adult insects appear as dark objects with a linear (bar-shaped) structure. Attempts to detect them by applying bar-shaped templates led ultimately to a linear feature detector which had a

good sensitivity for a reasonable range of insect sizes. The images from Experiment (c), adult *O. surinamensis* in wheat, were used to provide early test results using a basic but highly effective design of this type.

Larval insects are almost translucent, and the resulting severe lack of contrast prevents them from being detected reliably using either the thresholding or the linear feature detector approaches. Instead an approach which involved tracking around dark object boundaries, thresholding these to single pixel width and analysing the resulting path was developed. This gave long curved boundaries for grains but short straight paths for larvae, hence permitting them to be distinguished by finding the difference between the path length and the Euclidean distance between path ends. The images from Experiment (d), larval *O. surinamensis* in wheat, were analysed by this method. In the light of subsequent work, an attempt was made to apply an improved linear feature detector to the data from Experiment (d).

Rodent droppings were found to be relatively dark and thus candidates for detection by the application of thresholding. However, simple global thresholding was found not to be sufficiently reliable, and was replaced first by adaptive thresholding and then by the hybrid adaptive threshold procedure. Experiments (e), rat droppings in wheat, and (f), mouse droppings in wheat, describe results obtained using an early form of this procedure.

The integrated hybrid adaptive thresholding and linear feature detector package was tested by application to the following sets of images:

- Experiment (g): three ergots in wheat
- Experiment (h): six species of adult beetle in wheat
- Experiment (j): wheat with permitted admixture (from first source)
- Experiment (k): wheat with permitted admixture (from second source)
- Experiment (l): wheat heavily infested with insects (from first source)
- Experiment (m): wheat heavily infested with insects (from second source)

In Experiment (g), the number of grains in each image far exceeded the final specification level of 25, with the result that the integrated detection system was subject to memory overflow. To provide a reasonable test of the methodology, the linear feature detector routine was switched off for this experiment.

Imaging in the NIR region

Materials used in the images

Insects used were *S. granarius*. Wheat used for culturing was variety Mercia. Infested wheat kernels contained large, probably final instar, larvae. These were identified to developmental stage by visual inspection of X-ray radiographs. Uninfested kernels for use as control samples were also obtained by this method at the same time and from the same culture. Immediately prior to imaging it was confirmed by visual inspection that the outer surface of each infested kernel remained intact, free from any signs of the commencement of insect emergence.

Sample presentation

Samples were presented to the camera as four alternating rows of five clean and five infested kernels (Figure 8). Kernels were positioned with the crease away from camera and such that the germ-end was oriented downwards in the captured image. Kernel positions were fixed by mounting the sample on transparent adhesive plastic sheet supported by a rigid card frame. A set of four different images was obtained per 20-kernel sample: two NIR images at 1202 nm and 1300 nm, together with reference images both in the visible and X-ray regions. For the NIR and visible imaging, white card was used as the sample background.

Method of image capture

NIR images were captured in a similar way to the visible images above, but replacing the CCD camera with a Hamamatsu C2400-03 NIR vidicon camera/C2400 series controller and Fujinon 25 mm lens, positioned approximately 15 cm above the sample. Lighting for NIR imaging was from two 100 W standard lightbulbs, reflected on to the sample *via* an aluminium foil-covered card cylinder (approximately 20 cm diameter x 20 cm height) to give indirect diffuse lighting only. Images at the wavelengths 1202 nm and 1300 nm were obtained by using the appropriate filter (bandpass between 8 and 10 nm, 25 mm diameter, image quality; Andover Corporation), fixed flush to the lens housing and edge-sealed (Blu-Tak). Images were both shade-corrected, to minimise the effects of non-uniform lighting, and contrast enhanced using the camera controller. To improve signal to noise ratio, recorded images were the average of 200 frames. Reference images in the visible region were recorded using the CCD camera, with the lens zoom set to give an image of the sample comparable to that obtained by NIR. X-ray reference images were recorded using a Hewlett-Packard Faxitron 43804N camera (2.5 min exposure at 12.5 keV) and importing the radiograph to the computer by digital flatbed scanner.

Images recorded

Several replicate samples were studied. The set of four images with the best image quality is presented in this report [set (n) in Table 2].

Methods of image analysis

The digitised images obtained were inspected visually for qualitative differences. Contrast enhancement and sharpening were used to aid inspection. Machine vision methods were not investigated in this experiment. In addition to the four different images recorded from the same sample, a fifth composite NIR image was produced by subtracting the image at 1300 nm from the image at 1202 nm (Adobe Photoshop).

Imaging in the X-ray region

Materials used in the images

Materials used were as for the NIR imaging above, with the exception that here kernels containing pre-emergent adult insects were investigated. Also, control wheat was not obtained from the insect cultures, but was taken from the supply used for the culturing and cleaned.

Sample presentation

Samples were presented to the camera as a touching monolayer of kernels (10 cm x 10 cm), again mounted on adhesive sheet.

Method of image capture

The X-ray method was as described above.

Images recorded

One image was recorded for each of 3 replicate samples at each level of infestation of 0 or 5 infested kernels, giving a set of 6 images total, set (p) in Table 2.

Methods of image analysis

Two methods were used for the analysis of the images. The first method examined the shapes of all the objects within the image. This was done by thresholding the images, finding the edges of the objects and counting the numbers of pixels within each object and within the corresponding edges. The second method involved thresholding the image and labelling all the objects as before but the final stage was to generate the convex hull of each object. (The

convex hull of an object is the slightly larger shape formed by filling all the holes and boundary indentations.)

RESULTS

Imaging in the visible region

Using the fifty images of 1 to 5 barley kernels in wheat [Experiment (a)], contrast enhanced and sharpened, 78% of the barley kernels were correctly picked out by eye by an independent operator but this took considerable care. It was concluded that the images were of sufficient quality for analysis by computer algorithms.

In the initial familiarisation experiment using the sixty images from Experiment (b) of rapeseed in wheat, minimisation achieved a noticeable improvement in threshold estimation: Figure 4c (with minimisation) shows fewer incorrectly classified pixels than Figure 4d (without). The analysis by simple thresholding with minimisation showed that all 150 rapeseeds were detected. (Figure 4 is reproduced from the first annual interim report on this project; Chambers *et al.*, 1996.)

The results from the test of the linear feature detector on the images of adult *O. surinamensis* in wheat [Experiment (c)] are shown in Table 3. The results are quite impressive, there being no false positives, and only one missing insect over the whole sixty images (image 48). The insect was missed because it was in contact with another insect. This does not represent a practical failure of the approach. An example of an image containing adult *O. surinamensis* analysed by the linear feature detector is shown in Figure 5.

The results in the analysis of the images from Experiment (d), larval *O. surinamensis* in wheat, using the path tracking method, gave the results shown in Table 4. An example of an image containing larval *O. surinamensis* analysed by this method is shown in Figure 6. In the sixty images, containing a total of 150 largely translucent larvae, there was a total of 16 false negatives but no false positives: to achieve an overall failure rate as low as 11% with such low contrast contaminants is highly encouraging. However, 9 of the false negatives arose because the larvae were too close to grains for the tracking mechanism to work properly, indicating that further improvements would have to await the development of a new approach to the problem. Using the improved linear feature detector, the number of false negatives fell to six and there was only one false positive.

The initial form of the hybrid adaptive threshold algorithm was tested rigorously to locate rat and mouse droppings. Example images containing rodent droppings, mouse droppings or ergot analysed by adaptive thresholding are shown in Figure 7. In Experiment (e), rat droppings in wheat, there were 17 apparent instances of false positives (Table 5), however, all but one arose from the analysis fragmenting single items of contaminant. Rodent droppings are prone to such fragmentation as they have a speckled appearance. There were 3 false negatives but two of these arose because of contacts between contaminants. Hence there was only one genuine false positive and one false negative, a creditable performance on the detection of 150 droppings. In Experiment (f), mouse droppings in wheat, there were 9 apparent false positives, all of which were due to fragmentation, and 4 false negatives, of which one was due to object linking (Table 6). To some extent the problem of fragmentation has been eliminated in later versions of the detector.

The results from applying the integrated hybrid adaptive thresholding and linear feature detector package to the images in Experiments (g) to (m) are shown in Tables 7 to 12.

In all six of these Experiments, the integrated package proved reliable: for example, in Experiment (k) in which the images contained permitted admixture, none was reported as having contaminants (Table 10). In the other five Experiments, the number of cases where the system either missed insects or made errors of interpretation was commendably few but each warrants attention for methods of achieving further improvement.

In Experiment (g): three ergots, a total of 8 items of ergot (in images 41, 48, 51, 67, 68, 70, 74 and 79) was missed (Table 7). These false negatives were attributable to the threshold on area of ergot being set erroneously at 150 pixels. This is easily correctable, by reducing the area threshold to 90, a value which is still larger than the largest insect. The only other error occurred with image 42, in which a single large piece of ergot was reported as two because of a bright streak across it. While this was technically inaccurate, it did not affect the correct reporting of the presence of unacceptable contamination. The detection of all three types of ergot in this experiment can therefore be regarded as completely successful. It should also be noted that the algorithm was able to cope with over 300 overlapping grains in each of these images, this being well above the specification of 25 grains per image.

In Experiment (h): six species of adult beetle, one of the five grain weevils was missed in image 7 due to memory overflow caused by a technical problem which would not occur in the final commercial system (Table 8). In image 10, two of the five foreign grain beetles were missed through being too close to grain kernels.

In Experiment (j): permitted admixture from first source, there were four false positive identifications as insects (Table 9). The first two, in image 3, were actually short sections of grain boundary which were only slightly above the critical size for identification as an insect. The other two, in image 5, were due to a piece of chaff looking very like an insect and an unusually long black end of grain.

There were six similar false positive identifications as insects in Experiment (l): wheat heavily infested with insects from the first source (Table 11). Three of these were due to grain boundaries and three to pieces of chaff. A further false positive, in image 9, was caused by a program fault which has since been corrected. There were two false negatives, in images 11 and 30, both caused by an insect being too close to a grain and a further one in image 26.

In Experiment (m): wheat heavily infested with insects from second source, there were five false positive identifications as insects (Table 12). These were due to a piece of chaff (in image 8), a dark foreign object (in image 18), a dark forked piece of straw which was interpreted as two insects (in image 19), and a dark shrivelled grain (in image 22). More importantly, there were seven false negatives in which insects were missed. Three of these plus the false positive indication of ergot (in image 7) were due to memory overflow, which is correctable. One occurred in the top border (in image 11), two were small (in image 24) and one was only half emerged from its kernel (in image 30).

In summary, the 172 images in these six experiments, containing at least 5000 wheat grains, were interpreted correctly as containing a total of 143 insects and 60 pieces of ergot, apart from 39 named instances of false positives or false negatives. Of these false positives and false negatives, we have shown that we can immediately account for and eliminate 14 of them, leaving 25 instances out of 203. Seventeen of these are genuine false positives, and the operator's attention would be drawn to them, so the overall failure rate for insects is extremely low and that from ergot is negligible. If the maximum number of grains per image were restricted to 25 (which was not the case for any of these sets of images), the failure rate would plummet dramatically.

Imaging in the NIR region

Figure 8 shows the image of the sample of 10 infested and 10 control kernels in the visible region. There were no consistent differences with infestation, although a few of the infested kernels had small light areas on their surface. Several of the kernels had blackened germ-ends (pointing down the page), although in the bottom two rows these are slightly obscured by shadow.

Figure 9 shows the same sample, imaged in the NIR region at 1202 nm. The control kernels are uniformly dark, except at both their ends which are lighter. In addition, the blackened germ-ends are no longer detected. These observations indicate that there are significant differences between those images recorded in the visible and NIR regions and confirm the effectiveness of the filter. Importantly, the infested kernels appear more extensively and more consistently patchy at 1202 nm than in the visible.

Figure 10 shows the sample imaged at 1300 nm. The exposure for this image was not optimal because the settings used to record at 1202 nm were deliberately retained unchanged. Figure 10 looks qualitatively very similar to Figure 9. The image shown in Figure 11 is the difference between the two NIR images (1202 nm - 1300 nm). The quality of this image is poorer than those from which it was derived because the difference signal is smaller and therefore less easily distinguished from background noise. Nevertheless, it is clear that whereas the control kernels remain of fairly uniform shade, the infested kernels appear even more patchy and distinct from the controls. In this combined image, one of the 10 infested kernels is indistinguishable from the uninfested kernels (top row of infested, right-most kernel). This may have been caused by the lighting of the right-hand-most column of kernels being too poor for successful imaging (see Figure 10 especially).

Figure 12 shows the corresponding image recorded in the X-ray region. Although it appears that internal infestation can be detected well using the (1202 nm - 1300 nm) image, the patches present on the infested kernels of Figure 11 do not coincide with the cavities of the infested kernels of Figure 12.

Imaging in the X-ray region

In the first method, based on object shape, the optimum threshold on the difference between the number of pixels in the object and the number in the perimeter was found to be 14. This threshold was determined by adjusting the threshold manually for minimum classification error, there being insufficient data for a fully automatic procedure (though in fact the setting appeared to be non-critical). An example of this procedure is given in Figure 13. The method found all infested kernels (respectively one, two and one in number) in typical portions of respective sizes 100 x 100, 80 x 80, 80 x 80 pixels taken from the three images which included infested kernels. However, it also found a number of false alarms. Some were predictable, around the boundaries of the image, where the grains were only partially visible. In a practical method, grains would be kept entirely within the boundary. There were other false alarms: for example only a small part of a particular grain was light enough to be thresholded and one grain had a crease which gave an uneven outline when thresholded.

The convex hull procedure, designed to detect the dark area within the kernel which corresponds to the insect feeding cavity, also appeared to work fairly well. However, in all three images including infested kernels there were cases where the method produced an over-convex shape and included excessive amounts of equally dark background. Usually, although not always, this occurred where two or more grains were close or touching, as in the example given in Figure 14.

DISCUSSION

Imaging in the visible region

The nature of the inspection process

Automated visual inspection is part of the larger subject of image analysis, which is concerned with the processing and extraction of information from visual scenes. Such scenes originating from industrial applications are subject to special problems, not least the fact that the final application must be implemented in real time. If industrial processes are to be monitored or controlled, any image processing algorithms must keep up with the flow of information from products on a production line. Other than time, relevant parameters are cost, reliability and robustness. All these considerations were important in the present project to design an inspection system for detecting contaminants in grain.

The inspection task consists of four stages. First, images must be acquired with well-resolved, undistorted detail. Pre-processing operations can help restore images to a pristine state but these cannot be totally effective, so weaknesses in acquisition must be eliminated as far as possible. These operations can slow down the final system or make hardware implementation more expensive.

The second stage is that of searching for potential contaminants. In general, searching is unconstrained over the whole of every image and so is expensive in computer processing. Search is most problematic with objects of unknown shape, size or nature. It would be useful if a grain inspection system could alert operators to the presence of *any* types of unacceptable artefact other than those of highest priority (insects, ergot and rodent droppings). Even objects of specific shapes and sizes may appear in a variety of orientations. This means that considerable efforts must be allotted to the design of the relevant search algorithms.

The third stage is that of recognising the contaminants that are located by the search process. This is desirable because establishing which contaminants are present can be useful in deciding the subsequent course of action. For example, the presence of insect infestation could result in increasing contamination with time. Additionally, because the search process has been designed to be very fast, care in the design of the recognition stage can improve the overall reliability rate with fewer false alarms (false positives) and fewer missed contaminants (false negatives).

The final stage is the rejection of contaminants. For grain inspection, final outcomes could be: logging of the located contaminants; storing of images of contaminants in case of queries and storing of the contaminants by use of a rejection mechanism.

Steps in the investigation

The grain inspection process is complex and the necessary algorithms are intricate to design and test. Our investigations have been progressing through a number of key steps:

1. Familiarisation with the grain inspection problem.
2. Consideration of the contaminants that could arise, and design of algorithms to detect them individually.
3. Establishing the relative importance of the various contaminants, the necessity for detecting and recognising each of them, and the required accuracy in these judgements. In

particular, the need to decide the proportions of false positives and false negatives which might be reasonable, and the expense of achieving these proportions.

4. Grouping various contaminants under suitable headings. Dealing separately with each contaminant would be prohibitively expensive.
5. Detailed design and testing of algorithms to search for and detect the various groups of contaminants.
6. Integration of the overall search system for different types of contaminant.
7. Design of the subsequent contaminant recognition system.
8. Improvement of the execution times of the various parts of the algorithm.
9. Testing the overall system and validating its effectiveness, in terms of both reliability and execution rate.
10. Conceptual design and specification of the final hardware system. This stage cannot be divorced from design of the overall software system, as many algorithms would not run well on commercially available types of processor.
11. Anticipation of the likely market and possible vendor or types of vendor for such a system.
12. Attending to intellectual property issues.
13. Helping a vendor overcome problems in development and achieve a commercial system.

We have completed the first ten steps such that attention can now be directed to completing the final three.

Ensuring relevance in the investigations

For the initial research, images were recorded of samples which had been prepared in the laboratory [Experiments (a) to (f), (h) and (n) to (p)]. Once it had been proven with these samples that the proposed approaches were successful, their relevance to practical situations was tested by including samples obtained from the trade [Experiments (g) and (j) to (m)]. The acquisition of authentic samples of grain with insect species other than the saw-toothed grain beetle used in the earlier investigations required the retuning of the algorithms. In addition, analysis of the authentic samples of permitted admixture demanded additional recognition capabilities.

Available techniques

At the start of the project, a range of techniques was available which could be of use for analysing the images and detecting contaminants. Perhaps the most obvious was thresholding to convert the original grey-scale images to binary, thereby immediately segmenting the objects from their backgrounds. However, when objects may be touching or partially occluding each

other, this technique becomes less satisfactory, and model-based vision approaches such as Hough transforms become more attractive and appropriate. Even when images are less crowded, thresholding can be problematic especially because of variable illumination, which may also be apparent over sequences of images, and at the very least thresholds need to be computed afresh for each image. Adaptive thresholding, in which the threshold values are automatically adjusted to match the local conditions in the image, partly overcomes these problems, though ultimately boundary pattern analysis based on edge detection needs to be invoked. There is also the possibility of region based analysis, involving thinning and analysis of the resulting skeletons. Next there are neural network approaches which are apparently effortless to set up since the computer merely has to be trained to recognise objects. However, this apparent ease of set-up – together with additional advantages such as adaptability to a range of instances – is offset by difficulties of how best to configure and train the neural networks, and in many cases such large neural networks would be required that impracticably large training sets would be necessary. Hence use of neural networks needs rigorous questioning in each individual instance. Next, there are possibilities for using genetic algorithms for designing detection templates. Finally, there are texture analysis methods which are apparently appropriate for this application, since groups of grain kernels have a distinctly textured appearance.

This great variety of methods provided an excellent possibility of solving all the problems of the project, though the variety in itself led to difficulties, since it was not at all obvious what the best strategies would be for arriving at a sound cost-effective real-time solution to the problems.

Selection of the most promising techniques

At the start of the project, all the above techniques were considered carefully and most were investigated. Texture analysis, neural networks and genetic algorithms were found to lack promise. At the same time, two especially promising lines of attack were starting to emerge: one was thresholding and the other was the linear feature detector. Thresholding has a basic capability for finding dark contaminants including rapeseeds, rodent droppings and ergot, with reasonable reliability; while the linear feature detector is intrinsically suitable for locating small adult insects with reasonable reliability [see Experiment (c), Table 3], though not insect larvae. More work on larvae brought useful results skeletonising object boundaries and tracking them to characterise object shapes and hence separate larvae from grains [see Experiment (d), Table

4]. Thus at this stage three main approaches had emerged, with thresholding hopefully being used in a number of places in the overall algorithm to lead to savings in computation.

The next step was to investigate neural networks to solve the remaining deficiencies. It emerged that neural networks would be unreliable and computationally inefficient if used on their own, but as a backup to the linear feature detector could improve performance significantly, and also lead to reliable detection of the larvae. This led to the abandonment of the boundary tracking approach, which had had some problems with skeleton spurs induced by image noise. More important, dropping a method which is only used for one detection task is beneficial as generic approaches with wider applicability help to minimise computational effort.

It was difficult to raise the reliability of the thresholding approach to the desired levels of accuracy, even after introducing adaptive thresholding techniques. For example, in an early test on rat droppings, there were problems in finding a suitable global threshold in six out of fifty images, and in the corresponding test on mouse droppings, there were problems in two out of fifty images. These problems did not disappear on applying adaptive thresholding, partly as each image was fairly uniformly lit, so that adaptive thresholding could not help substantially with the problem; and partly because the speckle in the droppings added to the difficulty of determining accurate thresholds. Accordingly a new approach, the hybrid adaptive threshold, was designed. This involved use of quite large median filters (well known filters for removing noise in images), and this approach was retained in one form or another until the end of the project. It provided the capability originally envisaged for the thresholding approach – that of being applicable to all dark objects, ranging from rapeseeds to droppings, ergot and large adult insects. This is clear from the results obtained using the final integrated inspection system in Experiments (g) to (m).

Although the use of neural networks in conjunction with the linear feature detector was highly effective, it involved additional computation and thus was a candidate for exclusion from the final algorithm. Further investigation showed that improved design of the linear feature detector could enhance its performance significantly, with the result that it was as good as the earlier combined linear feature detector with final neural network stage. Again, this strategy was adopted and retained in the final design.

Algorithms have to be highly effective and sufficiently powerful that they can be used in the generic detection of different types of contaminant to be suitable for inclusion in the final

design of the system. In these respects, the linear feature detector and hybrid adaptive filters were the best, whereas the neural network and boundary tracking approaches had to be abandoned, even though they had been successful algorithms in their own right. It was concluded that the most promising solution to the problem of detecting the required range of grain contaminants would be based on a package integrating hybrid adaptive thresholding and linear feature detectors.

Detection of rodent droppings and ergot

Early in the project it was found that rodent droppings and ergot could be located reasonably reliably by global thresholding operations. However, more detailed investigations showed that for highest reliability, some means of eliminating small objects had to be invoked. Accordingly we devised morphological techniques including specially tuned operations which either shrink objects or slightly expand them. Later on we found that the hybrid adaptive threshold technique gave further increases in performance, by virtue of its use of quite large median filters to suppress noise and small noisy objects including shadows between the grains. Part of the reason for this success was also due to the fact that the rodent droppings are often speckled, and the hybrid adaptive threshold was able to look for regions of relative darkness, ignoring the speckle. In general we found that ergot is not subject to speckle (though it sometimes contains bright shiny streaks), and is often somewhat darker, longer and thinner: we were able to use these facts to help distinguish it from rodent droppings, albeit with some degree of uncertainty. Early results for rat and mouse droppings were each in the region of one false positive and one false negative per 150 instances of droppings. For ergot the results were better - no genuine errors in a trial of 60 instances of ergot in a complex background of wheat grains.

Detection of adult and larval insects

Adult insects appear as dark objects against the light wheat background. They also have a linear (bar-shaped) structure and we have been able to evolve rigorous procedures for detecting them by searching for this shape, using our so-called linear feature detector. By careful design of templates used for detection (templates are idealised forms of relevant objects which are used for matching purposes in images) we can largely eliminate false alarms from the dark shadows between grains of wheat, and false alarms around grain boundaries. However, this proved more difficult with insect larvae (which also appear in a linear bar-shaped structure), and in that case the templates had to be tuned much more carefully for the specific sizes and shapes of the larvae. Even so, there is little possibility of detecting larvae which are

too close to grains. The underlying problem is that larvae are almost transparent and present extremely low contrast: they are only rendered visible when they appear on a plain light background between grains. The smallest adult insects give very few pixels within their boundary which attain the required degree of blackness to ease detection, and hence they can be mistaken for noise in some cases. This is more likely to occur in the immediate proximity of grain boundaries, where the linear feature detector masks partly overlap the grains and hence give lower signal-to-noise ratio (and hence reduced sensitivity of detection). Overall, these factors mean that false alarm detection is needed in some form.

To minimise false alarms, the number of grains per image needs to be strictly limited – this being extremely important for the linear feature detector, though relatively unimportant for the hybrid adaptive threshold approach. Specifically, we have targeted the linear feature detector on 256 x 256 pixel images containing not more than 25 grains, insisting that the grains should be in a monolayer and seldom touching or overlapping. Nevertheless, the algorithm is robust enough to cope with some grains which touch or partially overlap, or which cast shadows near each other. The evidence from Experiments (d), (e) and (f), where contact between objects led to difficulties in analysis, suggests strongly that if the number of grain kernels in each image is restricted, the performance will improve significantly.

The limit on the number of grains per image means that more images have to be handled in the available time. Thus the algorithm has to be fast. We have decided that it is more reliable and rigorous to design algorithms that run sufficiently fast than it is to devise algorithms which can take in more grains at once, and which necessarily have additional checking procedures which might then slow them down even more, perhaps not even giving such a reliable result in the end.

In the current work the false negative rate for detection of adult insects varied. For example, in Experiment (c), the false negative rate was just 1 out of 150 insects. In Experiments (h), (l) and (m) this rose to 3 out of 30, 3 out of 28 and 7 out of 88 insects respectively. This rise was probably due to increase in the variety in insect sizes and shapes, and the introduction of significant amounts of permitted admixture. On the other hand, the speed of the analysis by this time was very much higher than in the earlier Experiments. Excluding false negatives due to correctable technical problems and those due to proximity to other items causes the false negative rates to fall to 0 out of 150, 0 out of 30, 1 out of 28 and 3 out of 88 adults of various species in Experiments (c), (h), (l) and (m) respectively.

Detection of foreign seeds

In practice the importance of detecting foreign seeds in a commodity depends on the individual contract. In extreme cases wheat kernels, for example, may be undesirable in barley but more commonly such presence would not cause concern. On the other hand, the repeated presence of rapeseed in wheat might be of interest and require investigation of its source.

In this project, a limited investigation of the presence of barley in wheat showed that distinction would be difficult. As already described however, the exercise did confirm that the quality of the images produced by our capture system would probably be sufficient for subsequent analysis by machine. In another early investigation, we recorded images of rapeseed in wheat. We have now found that some recognition capability for rapeseed can usefully be retained as a by-product of other detection processes. In all cases we encountered, it appeared to be straightforward to detect rapeseeds with 100% reliability. However, rapeseeds if grouped together might be misinterpreted, for example as mouse droppings.

Estimation of reliability

Thus far in this report, errors have been reported as false positives and false negatives. These can be used to estimate reliability of detection and recognition. This is usually expressed as probabilities of variation from the average true value of contaminant items present. For example, if the average number of items is 5.63 in 1000 images, then the false positive and false negative rates might be expressed as probabilities each of 0.1 that the average number of items actually found would be less or greater than 5.63.

Since different items of contamination are detectable with different difficulties, the false positive and negative rates will vary between types of contaminant and even within one type. For example, large adult insects (detected by hybrid adaptive thresholding) will have different rates from those for smaller adult insects (detected by the linear feature detector).

Failure rates can really only be quoted with meaningful precision when the prototype apparatus has been tested with large numbers of real samples. In the meantime, the figures quoted in the suggested specification represent conservative estimates, on the basis of the studies reported here, and should be regarded more as starting points for further improvement than as the best ultimately achievable. In any case these conservative figures represent a very substantial improvement over existing examination methods based on subjective visual inspection.

Contaminant recognition

The detection of contaminants had to be followed by a final recognition stage for two reasons: to make the system more useful and to improve confidence in the results by providing positive identification rather than a statement that a contaminant has been found. The latter reason is important as the contamination detection algorithms, including both the linear feature detector and the hybrid adaptive thresholding system retained in the final design, are very fast and intended to reveal all faults while minimising false alarms. Without a system to eliminate a small number of false positives, the main detectors could only be tuned for reduced detection capability, effectively equalising the numbers of false positives and false negatives. Our approach was to eliminate most of the false negatives and then identify and eliminate as many false positives as possible. This approach was introduced at the final system integration stage when the overall timing of the system was also being optimised.

The task of the recognition system was to consider all the regions labelled by the two main detectors and to analyse their shapes and proximities to grains. There would not be time to perform this task rigorously, nor would there be the need, as the main part of the detection and identification process had already been carried out. Thus the recognition system only had to distinguish the following categories:

- ergot
- rodent dropping
- insect
- other non-grain object

In practice, the system would be able to provide more information than this. For example, large and small adult insects could be distinguished (primarily on the basis of the area of the dark contaminant) and while this information might be of only occasional value, it was provided by the system with little effort and was retained. Similarly, the "other non-grain object" category was easily subdivided into "shadow between grains", "black end of grain", "rapeseed" (all of which are useful diagnoses as these categories could otherwise be counted as false positives for some types of insect) and "large object". The latter relates to the possibility of a large dark object which could take a lot of time for the system to analyse and should be logged because it is undesirable, though there is no point in examining its nature in detail by computer. As a one-off contaminant, the system could have no knowledge of what it is, and a human operator would need to be summoned for identification. On the other hand, subdivision

of the "other non-grain object" category by the recognition system would, depending on contract and therefore system settings, allow at least some of the labour in distinguishing acceptable objects from the unacceptable to be undertaken automatically.

The rarity of a type of contaminant is a potential problem. By far the majority of images would have no contaminant. The most frequent contaminants would probably be insects, rapeseeds, droppings or ergot. The rarity of other types of contaminant means that there is little information for training, so further categories are difficult to include. Some of the present ignorance of the nature and relative frequency of the more unusual types of contaminant may be reduced when the system is tested in trials with substantial consignments of grain.

Automatic threshold setting and updating

Throughout this report there are references to thresholding, which is a vital procedure, especially when searching for dark contaminants such as ergot. It involves converting the image to a binary (black and white) form in which the black regions are intended to denote dark contaminants. Unfortunately, applying a global threshold can be unreliable, and in that case adaptive thresholding can be useful. However, even when thresholding is a valid procedure, the threshold value still has to be determined reliably, and this can be a non-trivial process. We have tested several methods for determining thresholds, notably the entropy method and its derivatives, and histogram analysis. Both have appeared suitable, though with our currently available images the latter is slightly better. A final decision on means of setting and updating thresholds will have to wait until large numbers of images are being gathered automatically in practical situations. This is because the task has to be optimised for a valid sequence of images, and cannot be done rigorously on an individual image basis.

Improvements in the speed of the algorithms

There are two crucial stages in the later development of algorithms for contaminant detection: one is the integration of algorithms into a complete system, the other is the improvement in speed to allow the integrated system to perform in real time. It is likely that if the initial working versions of the various algorithms had been linked together to form a complete system, the total execution time would probably have been in the region of 30 seconds per 256 x 256 pixel image. (This assumes that the Pentium processor running at 200 MHz on which we have been developing the algorithms is being used for timing.) Calculations show that, to achieve the target 3 kg in 3 minutes using images of no more than 25 kernels each of 50 mg, some 2400 images would have to be examined in the 3 minutes, and each image would have to

be processed in just 0.075 seconds. Thus a speed-up factor of about 400 was required. Very little has been published on the design of real-time systems for the inspection of food products which might suggest how to proceed in the present case (but see Davies *et al.*, 1995). The usual way of making substantial improvements in speed is to institute manageable improvements by software alone, then build hardware systems to achieve the remaining improvements, using well-known parallel processing principles (i.e. using a number of processors, often of a relatively cheap variety so that greater numbers of them and hence larger speed-up factors can be afforded).

We obtained a handsome speed-up by substantial recoding of the algorithms. After eliminating the more obvious time consuming sub-routine calls, and thinking carefully about accessing and ordering of imaging functions, the speed had been improved by a factor of two or three. To achieve more, very great care was needed over the precise coding of the linear feature detector, the hybrid adaptive thresholding algorithm, and the object labelling routine. This last routine is a crucial aspect of the recognition process, which can slow the whole contaminant detection algorithm by a substantial factor. Two methods were adopted to speed this routine to realistic levels: (1) it was recoded as an abstract list-processing algorithm rather than as a conventional image propagation algorithm; (2) it was arranged that images in which large areas of contaminant have to be labelled and analysed would be labelled more crudely, and the operator's attention drawn to the fact that some substantial foreign body is present. Thus the category "large object" was added. These methods reduced the overall execution of the object labelling time down to around 0.2 seconds in a worst case, and close to 0.0 seconds in the absence of any contaminants.

The linear feature detector was also speeded up significantly, by arranging a hierarchy of skimming masks and operations which would progressively eliminate increasing amounts of the image from consideration. This was highly successful and resulted in the overall execution time for this function of less than 0.1 seconds – nearly 80 times faster than the original time of 7.6 seconds for a typical image containing five insects.

Finally, the hybrid adaptive thresholding algorithm was speeded up very significantly, from around 8 seconds to less than 0.1 seconds, by applying various procedures to the median filtering operation, and economising on the final location accuracy since this is not an important parameter. This proved sufficient to the task as the rat droppings and ergot which were particularly targeted by the hybrid adaptive threshold have significant size.

As a result of these measures, the current form of the integrated contaminant detection system basically runs in around 0.18 seconds. Some other operations remain to be speeded up but these will be intimately connected with the hardware environment provided as standard by the vendor of the system. Additional functions, including object labelling, depend on the number of contaminants present. It is the average time for dealing with contaminants that matters over the available three minutes given for each consignment. With the expected low numbers of contaminants and admixture it is expected that the mean additional execution time per image would be as low as 0.02 seconds. Second, for consignments with substantial contamination, it would be justifiable to work on the basis of finding the first 50 images containing significant numbers of contaminants, and not examining the remainder.

In practice the total execution time per image can be reckoned at around 0.2 seconds for images containing an average of 25 kernels. The advent of faster computers suggests that their application to the system should bring further reduction in time down to the target of 0.07 seconds per image. The processor used for our assessments is a 200 MHz Pentium PC, bought in January 1997. Already similar machines can be bought with a 300 MHz clock rate, and we anticipate that by the time our system is on the market around July 1998, relevant processing rates for PC's in the £2000 bracket will be 400 MHz. This would mean that a complete real-time hardware processing system for analysing the images could be bought for less than the cost of two PCs, namely around £4000. Furthermore, the framestore facility provided by the vendor could well incorporate functions such as median filtering, morphological processing and/or convolution processing, any or all of which could be used to improve the processing rate and further reduce cost. This means that careful choice of vendor will be required if the most cost-effective machine is to be marketed. What we have achieved within this project is to show that even if no hardware functions are available on the vendor's hardware, a real-time system that can inspect grain for all the crucial contaminants can be produced for little more than the cost of two PC's and a framestore and camera. This assessment excludes mechanical mechanisms required for grain handling: such mechanisms will have exacting requirements but they are not the immediate business of this project as they will embody known technology.

Comparison of likely performance with existing devices

The results presented in this report show that with some further development a device based on the principles described will deliver the performance required by the UK trade. There are two types of existing device which perform tasks related to the detection of contamination

investigated here. The Tecator 310 GrainCheck is capable of detecting ergot and would be expected to detect rodent droppings (although perhaps not as effectively as our hybrid adaptive threshold detector). The detection of insects has not been attempted using this device and it is not certain that it would work. The GrainCheck has no ability to detect internal insects. It is able to sort and weigh the different fractions found within the sample, presenting the results as weight percent, for example. However, the scan rate of this device is slow such that analysis of the required amount of grain (3 kg) would take about an hour rather than the few minutes available for example at intake laboratories.

Optical sorters are designed primarily to remove undesirable material which is darker than the commodity. They are used routinely for the whole-product sorting of relatively high-value commodities such as top quality rice, pulses and nuts. NIR optics can also be used (for example to sort hulled from de-hulled oats; ESM, personal communication). The commodity is inspected as it falls through the air from a series of parallel single-kernel chutes. The crucial difference between this and machine-vision is that colour-sorters cannot obtain and use information on size and shape. Hence they cannot be made selective but will detect and remove any object which is darker than or which contains surface features darker than the acceptable level of discolouration, providing it is of similar particle size to that of the commodity. As adult beetles, rodent droppings and ergot are often darker than the grain substrate, colour-sorters could in theory remove such contaminants for confirmation by visual inspection. Sorting of ergot from rye has been reported (ESM, personal communication) and it is expected that sorting of rodent droppings might be possible. However, the colour-sorting of adult beetles is not possible because the optical differences are of insufficient magnitude for reliable detection and insects are too small and of insufficient weight to flow through a colour-sorter in a predictable manner. Colour sorting of insect larvae would be impossible for similar reasons. Colour-sorters would not be able to discriminate between these dark contaminants and the less important dark contaminants such as rapeseeds and kernels with black point, for example. Colour sorters do not classify contamination type.

For these reasons, it is believed that no existing device has a performance which is comparable to that expected from the work presented in this report. Once it had become clear that it should be possible to meet the needs of the trade for a contamination detector, it was possible to compile a specification for the system. This is presented in Appendix B.

Imaging in the NIR region

The detection of hidden infestation, i.e. kernels infested internally with insects, is a very challenging problem. Previous work showed that wheat kernels internally-infested with grain weevil larvae can be discriminated from clean kernels using just two NIR wavelengths (Chambers and Ridgway, 1996), despite being invisible to the naked eye. The value of the variable [log 1/R(1194 nm)-log 1/R(1304 nm); where R is the light reflected back by the sample] was found to be consistently lower with infestation. This suggested that including a simple NIR imager in the machine vision system would extend detection to internal insects. To investigate this, we studied the use of an NIR vidicon camera and filters at 1202 nm and 1300 nm.

Differences were seen between the infested and the control kernels when imaged at 1202 nm. This phenomenon, the appearance of light patches with infestation when imaging at only one carefully chosen NIR wavelength, was predicted from our previous work and is thought to be due to increased light scattering from the insect/insect cavity surfaces (Chambers and Ridgway, 1996). These differences were more obvious and more distinct from uninfested controls in the image obtained by subtracting that at 1300 nm from that at 1202 nm. This subtraction was expected to remove most of the contribution from physical light scattering, such that the composite image would map certain changes in chemical composition. Starch absorbs at around 1200 nm and it was thought that loss of starch due to insect feeding would be detected. However, the similarity of Figures 11 and 9 suggests that physical scatter remains in the subtracted image, which is therefore not mapping purely chemical differences. Comparison with the image of the same sample recorded in the X-ray region suggested that the NIR effect is not detecting the insect/insect cavity directly but rather changes in kernel composition as a result of the infestation (due to e.g. feeding, excreting). However, larvae might have been moving in the cavity as a result of heat generated from the lighting and this may have obscured detail in the NIR images.

Imaging in the X-ray region

It was hoped that using the first method, examination of object shape, would permit finding an internally infested kernel where the outline would be disrupted giving a small difference between the number of pixels in the object and the number in the perimeter. Although the method found infested kernels, the false alarms it gave, and the false negatives when the

thresholded infested kernel had a normal shape and therefore was not found, meant that the method was not sufficiently robust, so an alternative was investigated.

The second method used convex hulls generated from the objects so that the dark areas representing the empty cavity within an infested grain would be interpreted as part of the kernel and would be included when computing the intensity histograms and intensity gradients. This was imperative since the cavities are the characteristic areas of intensity which distinguish infested from uninfested kernels. The inclusion of background intensities within almost all the convex hulls meant that a sizeable number of dark pixels was seen in the intensity histograms of almost all of the objects in the image, making it impossible to distinguish them from the infested kernels by examining the histograms alone. It was concluded that the method would have to be much more sensitive and computation intensive to be capable of locating infested kernels reliably.

Needs for future research and development

- Add capability to detect external larval insects to the integrated hybrid adaptive thresholding and linear feature detector package
- Optimise thresholds and other adjustable parameters
- Use information from incorrect identifications to improve diagnosis
- Investigate use of high frequency fluorescent tube and framestore to improve image quality by presenting shapes which are more accurate
- Use prototype to confirm that grain variability, for example with variety, causes no problems
- Confirm performance with other cereal species including barley and oats
- Conduct further NIR investigations to test unknown samples and developmental stages other than the final instar larvae tested here, establish its accuracy and discover the origin of the response, so that performance can be optimised

CONCLUSIONS

1. A low-cost image capture system using readily available components has been developed and constructed: it provides monochrome digital images of sufficient quality for analysis by machine vision. The best method of providing even illumination and minimum shadow was found to be by using a circular fluorescent tube.

2. A wide range of techniques for the analysis of images and detecting contaminants was considered carefully and most were investigated. Two especially promising approaches were those based on thresholding and linear feature detection. Each needed substantial innovative modification to deliver the required performance for the present application.

3. Application of these approaches to images recorded in the visible region has shown that it will be possible to meet the primary trade need for a system which will detect adult and larval insects external to grain kernels, ergot and rodent droppings. This was demonstrated by testing samples generated in the laboratory, then including samples obtained from the trade to confirm relevance to practical situations.

4. Early results for rat and mouse droppings were each in the region of one false positive (false alarm) and one false negative (object missed) per 150 instances of droppings. For ergot there were no genuine errors in a trial of 60 instances of ergot.

5. Excluding false negatives due to correctable technical problems and those due to close proximity to other objects, which would be avoided by reducing the number of grains per image to 25, the false negative rates for the detection of adult insects in four experiments were 0 out of 150, 0 out of 30, 1 out of 28 and 3 out of 88 instances of various species of grain beetle pest.

6. Despite larval insects being almost transparent and therefore presenting extremely low contrast, the improved linear feature detector gave only 6 false negatives out of 150 larval *O. surinamensis*.

7. In practice, the system would be able to provide more information than this: large and small adult insects could be distinguished, and other objects recognised as not being grain

were easily subdivided into "shadow between grains", "black end of grain", "rapeseed", and "large dark object". For example, rapeseeds were detected with complete reliability.

8. Some of the present ignorance of the nature and relative frequency of the more unusual types of non-grain items may be reduced when the system is tested in trials with substantial consignments of grain.

9. To minimise false alarms, the number of grains per image needs to be limited. We suggest that each image contains not more than 25 grains, these being in a monolayer and seldom touching or overlapping. The total execution time of the image analysis and recognition package was reduced from about 30 seconds to around 0.2 seconds for such images. The advent of faster computers suggests that their application to the system should bring further reduction in time down to the target of 0.07 seconds per image. This would allow the system to handle samples at the required rate of 3 kg in 3 minutes.

10. The cost of the hardware to run the complete machine vision package, excluding that required for sample handling, is likely to be around £4000.

11. No existing device has a performance which is comparable to that likely by the system developed from the work presented in this report.

12. Once it was clear that the system should meet the needs of the trade for a contamination detector, a suggested specification was written. The performance figures it contains are conservative estimates, on the basis of the studies reported here, and should be regarded as starting points for further improvement, not the best ultimately achievable.

13. Attention can now be directed to protection of the intellectual property upon which this system depends, identification of a suitable vendor, and work with vendor to bring the system to market.

14. Grains which are internally infested with grain weevil larvae cannot be detected by imaging in the visible region because they show no consistent differences from uninfested kernels. However, such infested kernels were detectable from images recorded at carefully chosen wavelengths in the NIR region. The NIR effect appears to be detecting changes in kernel composition. Incorporation of an NIR capability into the inspection

system should permit the reliable and rapid detection of internal infestation but this should be preceded by further research to allow its performance to be optimised.

15. Two methods applied to the analysis of X-ray images of grains internally infested with adult *S. granarius* gave encouraging results and showed that with further research and improved images, useful information could be extracted. It is less likely that this could be achieved at the speed needed to meet trade needs. Although detection of internal insects could probably be achieved by incorporating an X-ray camera into the system, this would increase complexity and operator skills required and could make the cost of the ultimate system high. It was concluded that such incorporation would not be worthwhile at this stage.

PUBLICITY

The achievements described in this report represent significant advances in intellectual property. This must be protected to encourage commercial exploitation of the system to be based on it. Nevertheless, the authors have been active in the promotion of their work and have taken care to do this without compromising the protection. Summaries of progress have been promoted to agricultural researchers (Chambers *et al.*, 1997a) and to the cereals trade (Chambers *et al.*, 1997b). Brief reference to it has been made at the 1997 Royal Show and in a national broadsheet newspaper (Spark, 1997). The latter has led to the Foreign Office commissioning a freelance writer to promote the work abroad. The work will also be promoted at an Anglo-French meeting on food quality (Chambers, 1997) and at an international meeting of NIR researchers and users (Ridgway and Chambers, 1997). Once the intellectual property has been protected, the work will be further promoted in trade journals and primary scientific journals of repute.

ACKNOWLEDGEMENTS

The authors are grateful to Robin Appel, Simon Hook, Paul Meakin and Chris Watson of the HGCA R&D Committee for guidance on the relative importance of contaminants and the supply of some realistic samples containing permitted admixture and insect infestation.

REFERENCES

Chambers, J. (1997) Revealing invertebrate pests in food In *Crop protection and food quality: meeting customer needs*, proceedings of BCPC/ANPP conference, 16-19 September 1997, Canterbury, Kent.

Chambers J. , Davies E. R., Ridgway C., Mason D. R. and Bateman M. W. (1997a). Detection of contaminants in cereals by image analysis poster at Image Analysis for Automating Agricultural Processes meeting, Society of Chemical Industry, 28 January 1997.

Chambers J. , Davies E. R., Ridgway C., Mason D. R. and Bateman M. W. (1997b). Detection of contaminants in cereals poster at Cereals 97, 11-12 June 1997.

Chambers J and Ridgway C (1996). Rapid detection of contaminants in cereals. In: *Near-infrared spectroscopy: the future waves*, eds Davies A M C and Williams P. NIR Publications, Chichester, UK, pp 484-489.

Chambers J. , Ridgway C., Davies E. R. and Patel, D. (1996). Rapid automated detection of insects and certain other contaminants in cereals. First annual interim report to the HGCA on project 0048/1/94, 12pp.

Chambers, J., Ridgway C., van Wyk, C.B., Baker, C.W. and Barnes, R.J. (1994). The potential of near-infrared spectroscopy for the rapid detection of pests in stored grain, Project Report 92, 51 pp., Home-Grown Cereals Authority, London.

Davies, E.R. (1993). Invited Viewpoint Sensing and Vision for food in the 1990s, Sensor Review, 13, no. 2, p. 3.

Davies, E.R. (1996). *Machine Vision: Theory, Algorithms, Practicalities*, Academic Press (2nd edition), pp. xxxi + 750.

Davies, E.R., Patel, D. and Johnstone, A.I.C. (1995). Crucial issues in the design of a real-time contaminant detection system for food products, Real-Time Imaging, 1, no. 6, pp. 397-407.

Hannah, I., Patel, D. and Davies, E.R. (1995). The use of variance and entropic thresholding methods for image segmentation, Pattern Recognition, 28, no. 8. pp. 1135–1143.

Keefe, P D. (1992). A dedicated wheat grain image analyser, Plant Varieties and Seeds, 5, 27-33.

Liao, K., Paulsen, M. R., and Reid, J. F. (1994). Real-time detection of colour and surface defects of maize kernels using machine vision, Journal of Agricultural Engineering Research, 59, 263-271.

Patel, D., Hannah, I. and Davies, E.R. (1993). Foreign object detection via texture recognition and a neural classifier, Visual Communication and Image Processing, SPIE, 2094, pp. 1291-1299.

Ridgway C. and Chambers J. (1997). Detection of insects inside wheat kernels by NIR imaging. Poster at Eighth International Conference on Near-infrared Spectroscopy, Essen, Germany, 15-19 September 1997.

Spark, D. (1997). War waged on the weevil Financial Times, 1 July 1997.

Zayas, I. Y., and Steele, J. L. (1990). Image analysis applications for grain science, Optics in Agriculture, SPIE, 1379, pp. 151-161.

APPENDIX A: GLOSSARY OF TERMS

For a source book on further information on image analysis and inspection issues see Davies (1996).

Adaptive thresholding	Thresholding in which the threshold value is automatically adjusted to match the local conditions in an image
Algorithm	Sequence of rules to conduct a mathematical process
Convex hull of a shape	The slightly larger shape formed by filling in any holes and boundary indentations
Double buffering	The provision of an additional storage buffer, so that incoming data can be written to store and output data read simultaneously. Useful for inspection applications, where the next frame can be grabbed without interruption while the previous frame is being read and processed
False negative	Instance of a failure to locate an object or contaminant
False positive	False alarm found when searching for an object or contaminant
Fragmentation	False positive arising from speckled or bright streaky appearance of a single contaminant item
Hybrid adaptive thresholding	An adaptive thresholding procedure which makes global and local tests of image data while computing the final binary image.
Linear feature detector	A detector which uses bar-like templates to detect objects or contaminants such as insects.
Median filter	Image processing filter which is scanned over the image to eliminate noise. It searches for isolated intensity values which are incompatible with neighbouring intensity values and which should be removed: in such cases they are replaced by more likely (median) values.
Pixel	Short form of 'PICture CELL'. The smallest resolved area in a digital image, normally taken to be square
Real time algorithm	A computer algorithm which is able to keep up with the normal flow of input data, as required, for example, on a production line.
Template	An idealised form of an object which is used for matching purposes in images.
Thresholding	The process of converting a grey-scale image into a binary image by applying a threshold: during conversion, any pixel which is darker than the threshold becomes black, and any other pixel becomes white

APPENDIX B: PRELIMINARY SPECIFICATION FOR AUTOMATED CEREAL CONTAMINANT DETECTOR SYSTEM DEVELOPED IN THIS PROJECT

System description

Kernel feeder: *input capacity of 3 kg batches to present to machine as 2400 portions each of about 1.25 g, including sieve to protect against blockage*
Image capture: *monochrome or linescan equivalent at 256 x 256 pixels, 8 bit precision*
Computer: *two off (or twin) Pentium 400 MHz, when available, or equivalent*
Framestore: *two frames of 256 x 256 pixels, 8 bit precision, with double buffered capability, possibly with linescan interface*
Output devices: *balance, VDU, printer, collection bucket, light and audible alarm*

Sample handling

Typical samples: *wheat, barley, oats*
Capacity: *batches of 3 kg in 3 minutes*
Dimensions of individual portions imaged: *6.5 x 6.5 cm*
Size of sample: *25 +/- 5 cereal kernels; no item to be less than 7 pixels from image boundary; distance between kernels to be not less than 2 x width of shadow between them*
Image capture rate: *13.3 images per second*:

Performance requirements

Types of contaminant: *droppings from rats and mice, ergot of different types, adult insects, larval insects external to kernels, rapeseeds, large objects*

Categories of contaminant distinguished:

(1) *rodent droppings/ergot**
(2) *adult /external larval insects*
(3) *rapeseeds*
(4) *large objects*
(5) *other contaminants*

Minimum limit of detection: *single items of each contaminant*

Mean false positive and false negative rates:

Category (1): *better than 0.3%*
Category (2): *better than 4%****
Category (3): *better than 0.2%*
Category (4): *better than 0.1%*
Category (5): *normally better than 10%***

---

* Droppings and samples of ergot of area less than 90 pixels (0.06 sq cm) may be classed as insects or, in some instances, rapeseeds, because of the similarity in shape.
** Will depend on shape, size, contrast, etc.
*** Quoted figures are for adult insects: small larval insects have lower contrast and the figures can only be guaranteed within a factor four of the quoted figures. For other details, see main text.

Response to permitted admixture: *does not report as contaminants permitted material such as hulls with or without grains; broken, shrivelled or squashed grains; grains with dark patches or black point; straw*

Presentation of results: *detection of contaminant of any type to result in illuminated light and audible alarm (optional), printout of weight of grain batch and number of each type of contaminant detected, any contaminants to be separated from main batch for subsequent inspection.*

Operation data
Image analysis procedure: *blob analysis, linear feature analysis, object labelling, object recognition, logging, interface to delivery/rejection/sample storage grain handling mechanisms*
Calibration mode: *preliminary screening of small samples of grain without contaminants; continual updating during operation*
Operating system: *Windows NT or other robust system for factory use*
Interface: *via controlling computer; possible printer output*
Operating temperature: *0 - 40°C (preferred less than 30°C to avoid electronic failures)*

Technical data
Power specifications: *for image analysis expected to be about 240 V, 2-3 kW*
Dimensions and weight: *for image analysis expected to be about 70 cm (h) x 30 cm (l) x 30 cm (w)*

All technical specifications will be subject to agreement between vendor and IPR holders, to ensure that variations (such as use of linescan camera, choice of framestore, alternative operating systems, etc.) do not interfere with the basic algorithm concept. It must be noted that the system developed by the IPR holders is a laboratory development system, which is not itself suitable for field use: thus the vendor would be *expected* to suggest his preferred adaptations to the specification laid down by the IPR holders.

Delivery to vendor: It is intended that the vendor will receive the following items under licence from the IPR holders:
1. Relevant know-how and experience
2. Relevant source code image analysis routines which pass image data and interpretation data via allocated memory areas.

In addition, the following will be available:
3. Further advice on how to handle unforeseen cases in the input images.

However, item 3 cannot be delivered without charge, as only items 1 and 2 have been paid for by the HGCA-funded project.

The IPR holders will *not* deliver executable code, as they will not be familiar with the vendor's marketable systems. In particular, *all responsibility for handling windows-type (image display) routines will rest on the vendor*: this reflects the 'proof of concept' approach adopted by the IPR holders.

NOTE: The above specification represents our present scientific and technological thinking. It has been written in good faith **but has not yet been vetted or approved by the RHUL External Relations Department or the CSL staff with similar responsibility**: hence no formal guarantees can be given over it. It is intended solely to clarify our present intentions with regard to development and marketing of the system.

Table 1: Description of authentic grain samples

Sample	Foreign objects present
Wheat with permitted admixture (first source)	wheat grains with hulls remaining; broken wheat grains; wheat grains with black point (germ-end blackened); wheat grains with dark patches; dark shrivelled wheat grain (one only); grain hulls; rogue seeds (various unidentified); foreign grain (barley); straw
Wheat with permitted admixture (second source)	As above plus squashed wheat grains; small wheat grains; dark, irregular dirt fragment (one only); black particle (seed case?; one only)
Insect-infested wheat (first source)	*C. ferrugineus* (approx. 200); *T. castaneum* (approx. 50); *A. advena* (approx. 30)
Insect-infested wheat (second source)	*S. granarius* (approx. 300); *O. surinamensis* (approx. 100); *C. ferrugineus* (approx. 20)

Table 2: Sets of images recorded

Experiment	Wavelength region	Description
a	Visible	Barley in wheat
b	Visible	Rapeseed in wheat
c	Visible	Adult *O. surinamensis* in wheat, kernels not touching
d	Visible	Larval *O. surinamensis* in wheat, kernels not touching
e	Visible	Rat droppings in wheat
f	Visible	Mouse droppings in wheat
g	Visible	Three ergots (two rye and one wheat) in wheat
h	Visible	Six species of adult beetle in wheat, kernels not touching
j	Visible	Wheat with permitted admixture (from first source), kernels not touching
k	Visible	Wheat with permitted admixture (from second source), kernels not touching
l	Visible	Wheat heavily infested with insects (from first source), kernels not touching
m	Visible	Wheat heavily infested with insects (from second source), kernels not touching
n	NIR	Wheat kernels internally infested with *S. granarius* larvae
p	X-ray	Wheat kernels internally infested with *S. granarius* adults

Table 3: Experiment (c), adult *O. surinamensis* in wheat. Actual number of insects in each image compared with number recognised by a linear feature detector.

Image number	No. of adults: found/ present	False positives	False negatives	Image number	No. of adults: found/ present	False positives	False negatives
1	0/0	0	0	31	3/3	0	0
2	0/0	0	0	32	3/3	0	0
3	0/0	0	0	33	3/3	0	0
4	0/0	0	0	34	3/3	0	0
5	0/0	0	0	35	3/3	0	0
6	0/0	0	0	36	3/3	0	0
7	0/0	0	0	37	3/3	0	0
8	0/0	0	0	38	3/3	0	0
9	0/0	0	0	39	3/3	0	0
10	0/0	0	0	40	3/3	0	0
11	1/1	0	0	41	4/4	0	0
12	1/1	0	0	42	4/4	0	0
13	1/1	0	0	43	4/4	0	0
14	1/1	0	0	44	4/4	0	0
15	1/1	0	0	45	4/4	0	0
16	1/1	0	0	46	4/4	0	0
17	1/1	0	0	47	4/4	0	0
18	1/1	0	0	48	3/4	0	1
19	1/1	0	0	49	4/4	0	0
20	1/1	0	0	50	4/4	0	0
21	2/2	0	0	51	5/5	0	0
22	2/2	0	0	52	5/5	0	0
23	2/2	0	0	53	5/5	0	0
24	2/2	0	0	54	5/5	0	0
25	2/2	0	0	55	5/5	0	0
26	2/2	0	0	56	5/5	0	0
27	2/2	0	0	57	5/5	0	0
28	2/2	0	0	58	5/5	0	0
29	2/2	0	0	59	5/5	0	0
30	2/2	0	0	60	5/5	0	0

Table 4: Experiment (d), larval *O. surinamensis* in wheat. Actual number of insects in each image compared with number recognised by a tracking method.

Image number	No. of larvae: found/ present	False positives	False negatives	Image number	No. of larvae: found/ present	False positives	False negatives
1	0/0	0	0	31	3/3	0	0
2	0/0	0	0	32	3/3	0	0
3	0/0	0	0	33	2/3	0	1
4	0/0	0	0	34	3/3	0	0
5	0/0	0	0	35	3/3	0	0
6	0/0	0	0	36	3/3	0	0
7	0/0	0	0	37	2/3	0	1
8	0/0	0	0	38	3/3	0	0
9	0/0	0	0	39	2/3	0	1
10	0/0	0	0	40	2/3	0	1
11	1/1	0	0	41	4/4	0	0
12	1/1	0	0	42	4/4	0	0
13	0/1	0	1	43	4/4	0	0
14	1/1	0	0	44	4/4	0	0
15	1/1	0	0	45	4/4	0	0
16	0/1	0	1	46	3/4	0	1
17	1/1	0	0	47	4/4	0	0
18	1/1	0	0	48	4/4	0	0
19	1/1	0	0	49	4/4	0	0
20	1/1	0	0	50	3/4	0	1
21	1/2	0	1	51	4/5	0	1
22	2/2	0	0	52	5/5	0	0
23	1/2	0	1	53	5/5	0	0
24	1/2	0	1	54	5/5	0	0
25	2/2	0	0	55	4/5	0	1
26	2/2	0	0	56	5/5	0	0
27	2/2	0	0	57	4/5	0	1
28	2/2	0	0	58	4/5	0	1
29	2/2	0	0	59	4/5	0	1
30	2/2	0	0	60	5/5	0	0

Table 5: Experiment (e), rat droppings in wheat. Actual number of droppings in each image compared with number recognised by a hybrid adaptive thresholding method.

Image number	No. of rat droppings: found/ present	False positives	False negatives	Image number	No. of rat droppings: found/ present	False positives	False negatives
1	0/0	0	0	31	3/3	0	0
2	0/0	0	0	32	3/3	0	0
3	0/0	0	0	33	3/3	0	0
4	0/0	0	0	34	6/3	3	0
5	0/0	0	0	35	3/3	0	0
6	0/0	0	0	36	7/3	4	0
7	0/0	0	0	37	4/3	1	0
8	0/0	0	0	38	3/3	0	0
9	0/0	0	0	39	3/3	0	0
10	0/0	0	0	40	3/3	0	0
11	2/1	1	0	41	4/4	0	0
12	1/1	0	0	42	4/4	0	0
13	1/1	0	0	43	4/4	0	0
14	1/1	0	0	44	5/4	1	0
15	1/1	0	0	45	4/4	0	0
16	1/1	0	0	46	4/4	0	0
17	1/1	0	0	47	5/4	1	0
18	1/1	0	0	48	4/4	0	0
19	2/1	1	0	49	5/4	1	0
20	1/1	0	0	50	5/4	1	0
21	2/2	0	0	51	5/5	0	0
22	2/2	0	0	52	5/5	0	0
23	2/2	0	0	53	4/5	0	1
24	2/2	0	0	54	5/5	0	0
25	2/2	0	0	55	4/5	0	1
26	2/2	0	0	56	5/5	0	0
27	1/2	0	1	57	5/5	0	0
28	2/2	0	0	58	5/5	0	0
29	2/2	0	0	59	8/5	3	0
30	2/2	0	0	60	5/5	0	0

Table 6: Experiment (f), mouse droppings in wheat. Actual number of droppings in each image compared with number recognised by a hybrid adaptive thresholding method.

Image number	No. of mouse droppings: found/ present	False positives	False negatives	Image number	No. of mouse droppings: found/ present	False positives	False negatives
1	0/0	0	0	31	3/3	0	0
2	0/0	0	0	32	3/3	0	0
3	0/0	0	0	33	3/3	0	0
4	0/0	0	0	34	4/3	1	0
5	0/0	0	0	35	3/3	0	0
6	0/0	0	0	36	3/3	0	0
7	0/0	0	0	37	4/3	1	0
8	0/0	0	0	38	2/3	0	1
9	0/0	0	0	39	3/3	0	0
10	0/0	0	0	40	3/3	0	0
11	1/1	0	0	41	5/4	1	0
12	1/1	0	0	42	4/4	0	0
13	1/1	0	0	43	3/4	0	1
14	1/1	0	0	44	5/4	1	0
15	1/1	0	0	45	4/4	0	0
16	1/1	0	0	46	4/4	0	0
17	2/1	1	0	47	4/4	0	0
18	1/1	0	0	48	4/4	0	0
19	1/1	0	0	49	4/4	0	0
20	1/1	0	0	50	4/4	0	0
21	3/2	1	0	51	3/5	0	2
22	2/2	0	0	52	5/5	0	0
23	2/2	0	0	53	5/5	0	0
24	2/2	0	0	54	5/5	0	0
25	2/2	0	0	55	5/5	0	0
26	2/2	0	0	56	6/5	1	0
27	2/2	0	0	57	6/5	1	0
28	2/2	0	0	58	6/5	1	0
29	2/2	0	0	59	5/5	0	0
30	2/2	0	0	60	5/5	0	0

Table 7: Experiment (g): three ergots in wheat. Actual number of ergot pieces in each image compared with number recognised by integrated hybrid adaptive thresholding and linear feature detector package

Image number	Ergot pieces: found/ present	False positives	False negatives	Image number	Ergot pieces: found/ present	False positives	False negatives
1	0/0	0	0	41	0/1	0	1
2	0/0	0	0	42	2/1	1	0
3	0/0	0	0	43	1/1	0	0
4	0/0	0	0	44	1/1	0	0
5	0/0	0	0	45	1/1	0	0
6	0/0	0	0	46	1/1	0	0
7	0/0	0	0	47	1/1	0	0
8	0/0	0	0	48	0/1	0	1
9	0/0	0	0	49	1/1	0	0
10	0/0	0	0	50	1/1	0	0
11	0/0	0	0	51	0/1	0	1
12	0/0	0	0	52	1/1	0	0
13	0/0	0	0	53	1/1	0	0
14	0/0	0	0	54	1/1	0	0
15	0/0	0	0	55	1/1	0	0
16	0/0	0	0	56	1/1	0	0
17	0/0	0	0	57	1/1	0	0
18	0/0	0	0	58	1/1	0	0
19	0/0	0	0	59	1/1	0	0
20	0/0	0	0	60	1/1	0	0
21	1/1	0	0	61	1/1	0	0
22	1/1	0	0	62	1/1	0	0
23	1/1	0	0	63	1/1	0	0
24	1/1	0	0	64	1/1	0	0
25	1/1	0	0	65	1/1	0	0
26	1/1	0	0	66	1/1	0	0
27	1/1	0	0	67	0/1	0	1
28	1/1	0	0	68	0/1	0	1
29	1/1	0	0	69	1/1	0	0
30	1/1	0	0	70	0/1	0	1
31	1/1	0	0	71	1/1	0	0
32	1/1	0	0	72	1/1	0	0
33	1/1	0	0	73	1/1	0	0
34	1/1	0	0	74	0/1	0	1
35	1/1	0	0	75	1/1	0	0
36	1/1	0	0	76	1/1	0	0
37	1/1	0	0	77	1/1	0	0
38	1/1	0	0	78	1/1	0	0
39	1/1	0	0	79	0/1	0	1
40	1/1	0	0	80	1/1	0	0

Table 8: Experiment (h): six species of adult beetle in wheat. Actual number of artefacts in each image compared with number recognised by integrated hybrid adaptive thresholding and linear feature detector package

Image number	Number of insects: found/present	False positives	False negatives
1	0/0	0	0
2	0/0	0	0
3	0/0	0	0
4	0/0	0	0
5	0/0	0	0
6	0/0	0	0
7	4/5 *S. granarius*	0	1*
8	5/5 *T. castaneum*	0	0
9	5/5 *R. dominica*	0	0
10	3/5 *A. advena*	0	2
11	5/5 *C. ferrugineus*	0	0
12	5/5 *O. surinamensis*	0	0

* Problem due to memory overflow (i.e. correctable)

Table 9: Experiment (j): wheat with permitted admixture (from first source). Actual number of artefacts in each image compared with number recognised by integrated hybrid adaptive thresholding and linear feature detector package

Image number	No. of artefacts: found/present	False positives	False negatives
1	0/0	0	0
2	0/0	0	0
3	2/0 (insects)	2	0
4	0/0	0	0
5	2/0 (insects)	2	0
6	0/0	0	0
7	0/0	0	0
8	0/0	0	0
9	0/0	0	0
10	0/0	0	0

Table 10: Experiment (k): wheat with permitted admixture (from second source). Actual number of artefacts in each image compared with number recognised by integrated hybrid adaptive thresholding and linear feature detector package

Image number	No. of artefacts: found/present	False positives	False negatives
1	0/0	0	0
2	0/0	0	0
3	0/0	0	0
4	0/0	0	0
5	0/0	0	0
6	0/0	0	0
7	0/0	0	0
8	0/0	0	0
9	0/0	0	0
10	0/0	0	0

Table 11: Experiment (l): wheat heavily infested with insects (from first source). Actual number of artefacts in each image compared with number recognised by integrated hybrid adaptive thresholding and linear feature detector package

Image number	No. of insects: found/present	False positives	False negatives
1	2/1	1	0
2	0/0	0	0
3	1/0	1	0
4	0/0	0	0
5	0/0	0	0
6	1/0	1	0
7	0/0	0	0
8	2/2	0	0
9	2/1	1*	0
10	0/0	0	0
11	0/1	0	1 (hooked on kernel)
12	1/1	0	0
13	0/0	0	0
14	0/0	0	0
15	3/3	0	0
16	0/0	0	0
17	0/0	0	0
18	0/0	0	0
19	1/1	0	0
20	0/0	0	0
21	0/0	0	0
22	0/0	0	0
23	0/0	0	0
24	0/0	0	0
25	1/1	0	0
26	1/2	0	1
27	1/1	0	0
28	5/3	2	0
29	7/6	1	0
30	4/5	0	1

* Problem due to fault in program, since corrected

Table 12: Experiment (m): wheat heavily infested with insects (from second source). Actual number of artefacts in each image compared with number recognised by integrated hybrid adaptive thresholding and linear feature detector package

Image number	No. of insects: found/present	False positives	False negatives
1	3/3	0	0
2	6/6	0	0
3	5/5	0	0
4	0/0	0	0
5	4/4	0	0
6	2/2	0	0
7	7/10 (+ 1/0 ergot)	1 (dark object)*	3 (dark object touching 2 insects)*
8	6/5	1	0
9	3/3	0	0
10	2/2	0	0
11	1/2	0	1
12	1/1	0	0
13	1/1	0	0
14	3/3	0	0
15	1/1	0	0
16	2/2	0	0
17	3/3	0	0
18	3/2	1	0
19	2/0	2	0
20	3/3	0	0
21	5/5	0	0
22	3/2	1 (dark shrivelled kernel)	0
23	2/2	0	0
24	9/11	0	2
25	1/1	0	0
26	2/2	0	0
27	3/3	0	0
28	1/1	0	0
29	2/2	0	0
30	1/2	0	1 (only half-emerged)

* Problem due to memory overflow (i.e. correctable)

LIST OF FIGURES

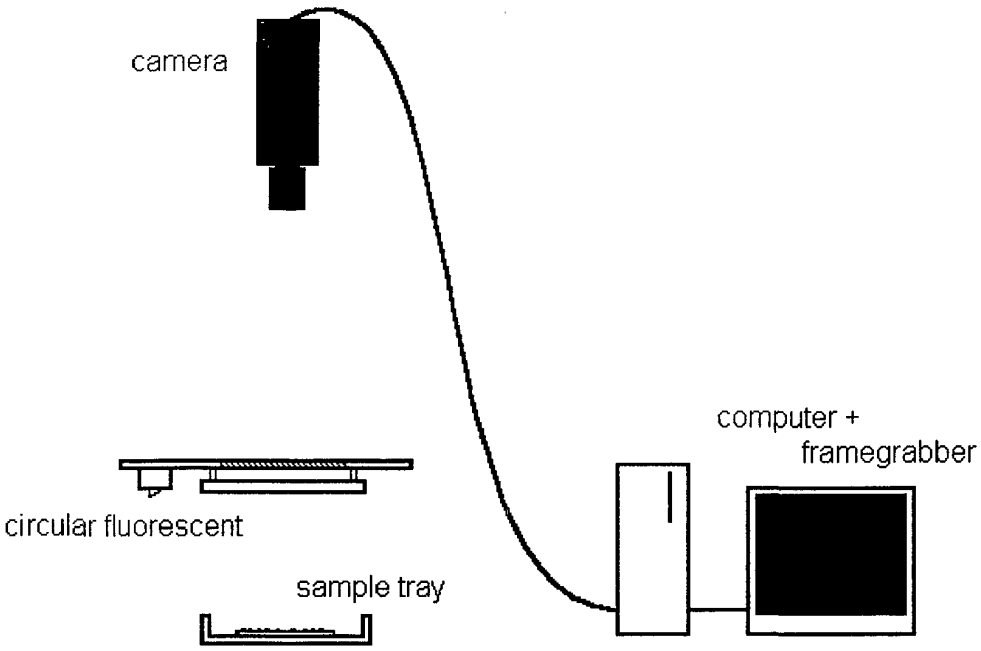

Figure 1: Image capture arrangement

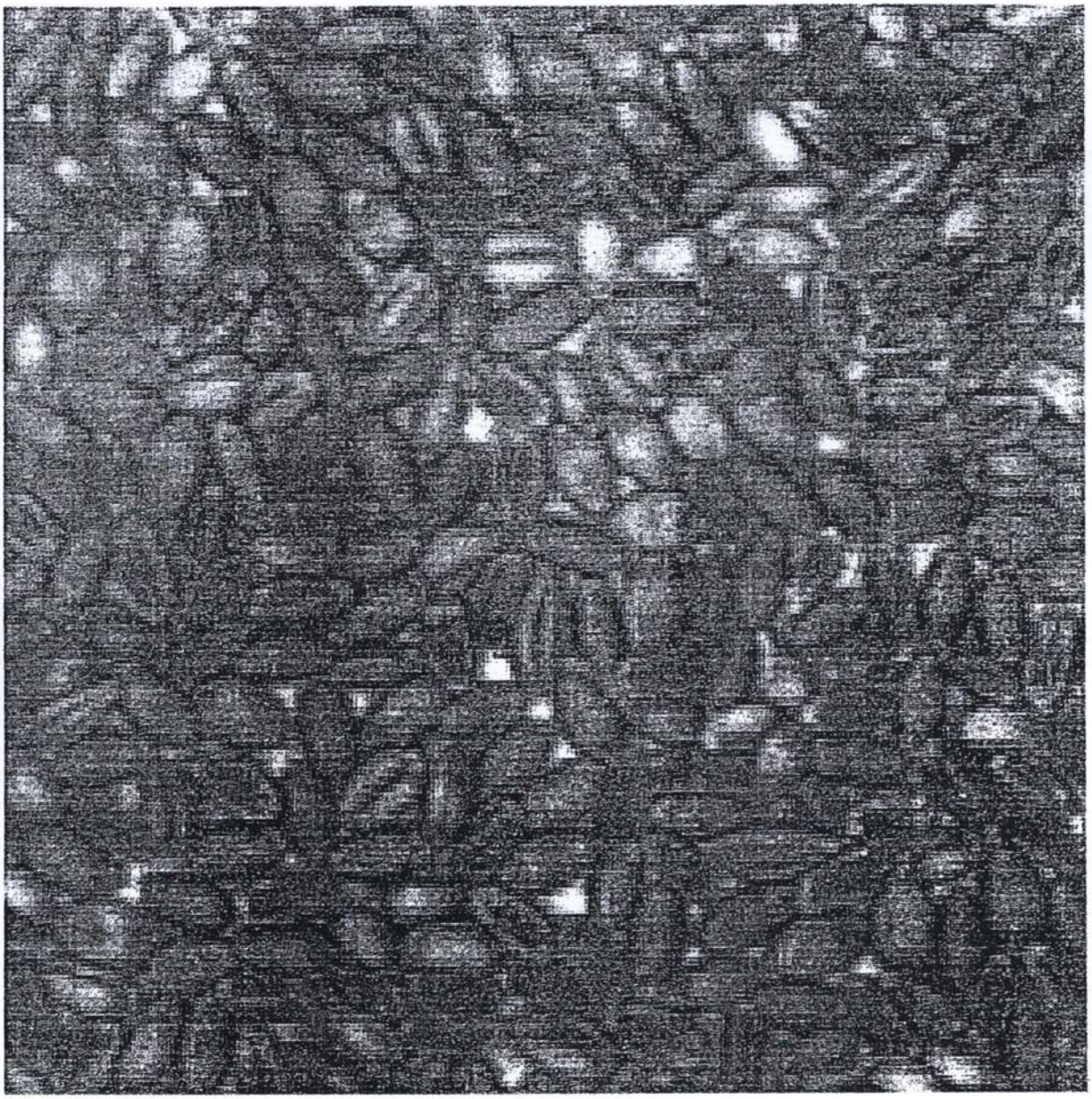

Figure 2: Experiment (a): example of image of wheat which was inspected by the naked eye for the presence of barley kernels

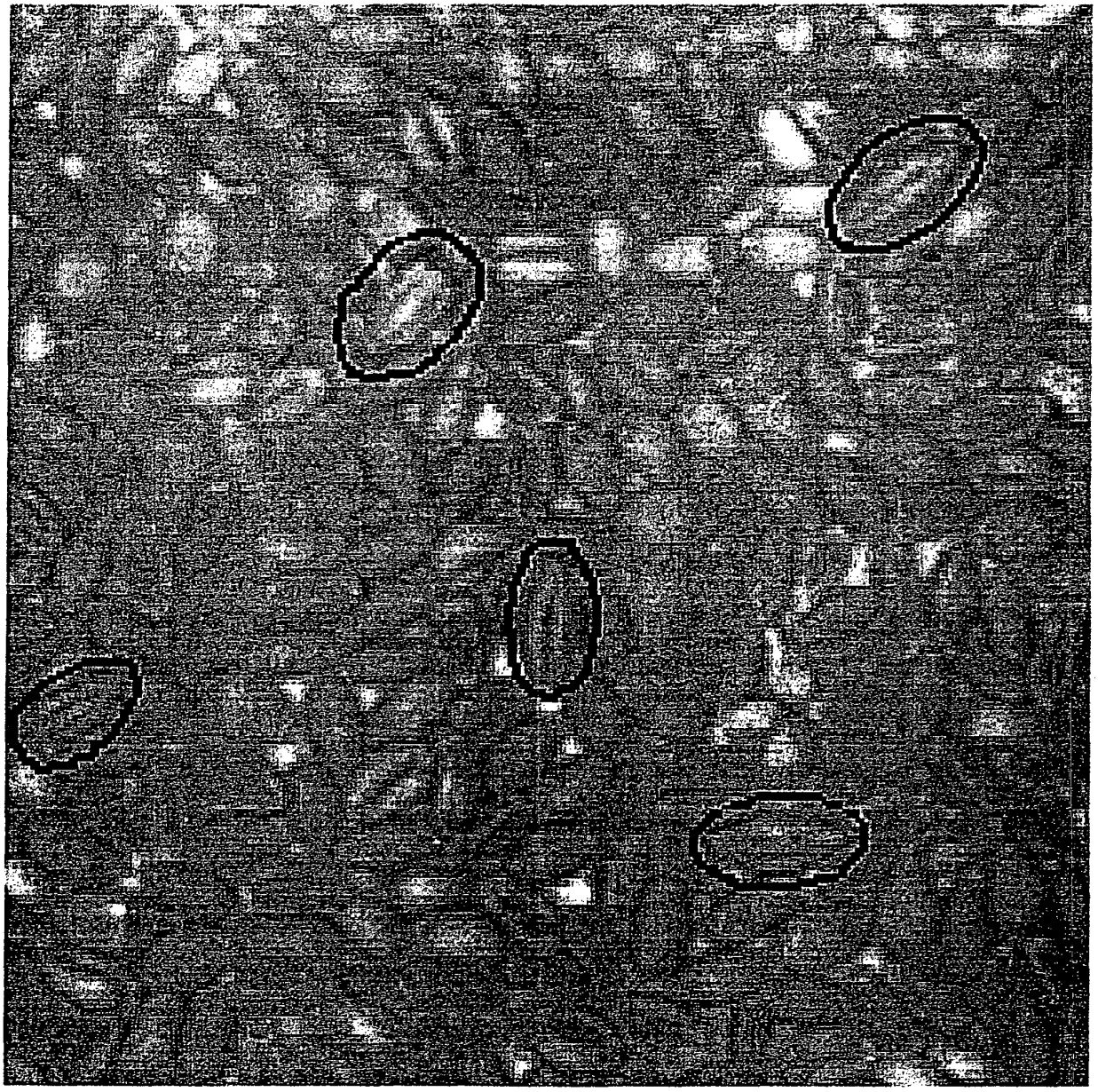

Figure 3: Experiment (a): image of wheat shown in Figure 2 which was marked for the presence of barley kernels by the person who prepared it

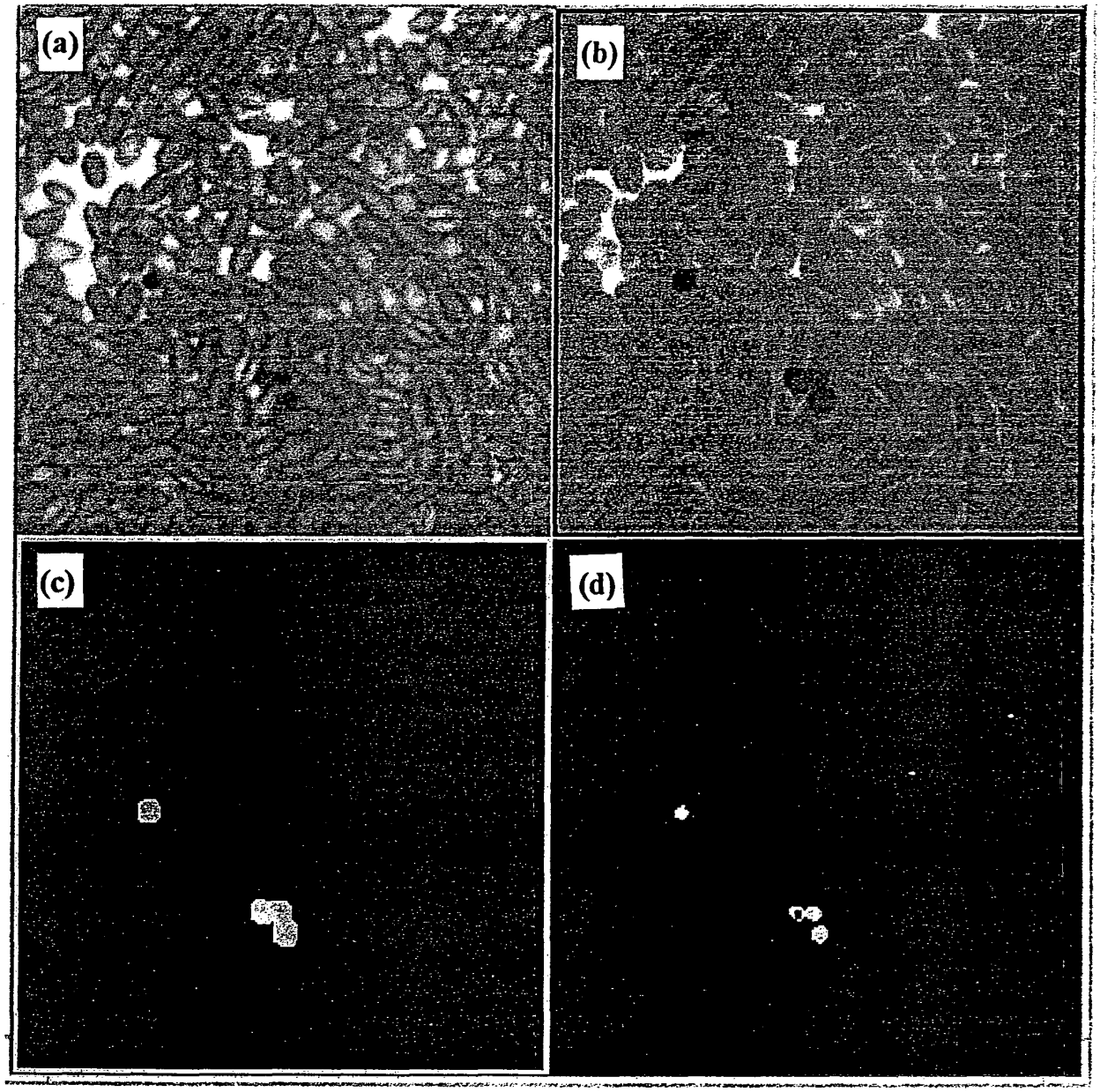

Figure 4: Experiment (b): image of four rapeseeds in wheat showing the benefit of including a minimisation pre-processing treatment: (i) original image, (ii) after minimisation, (iii) after minimisation and thresholding, and (iv) after thresholding but without minimisation

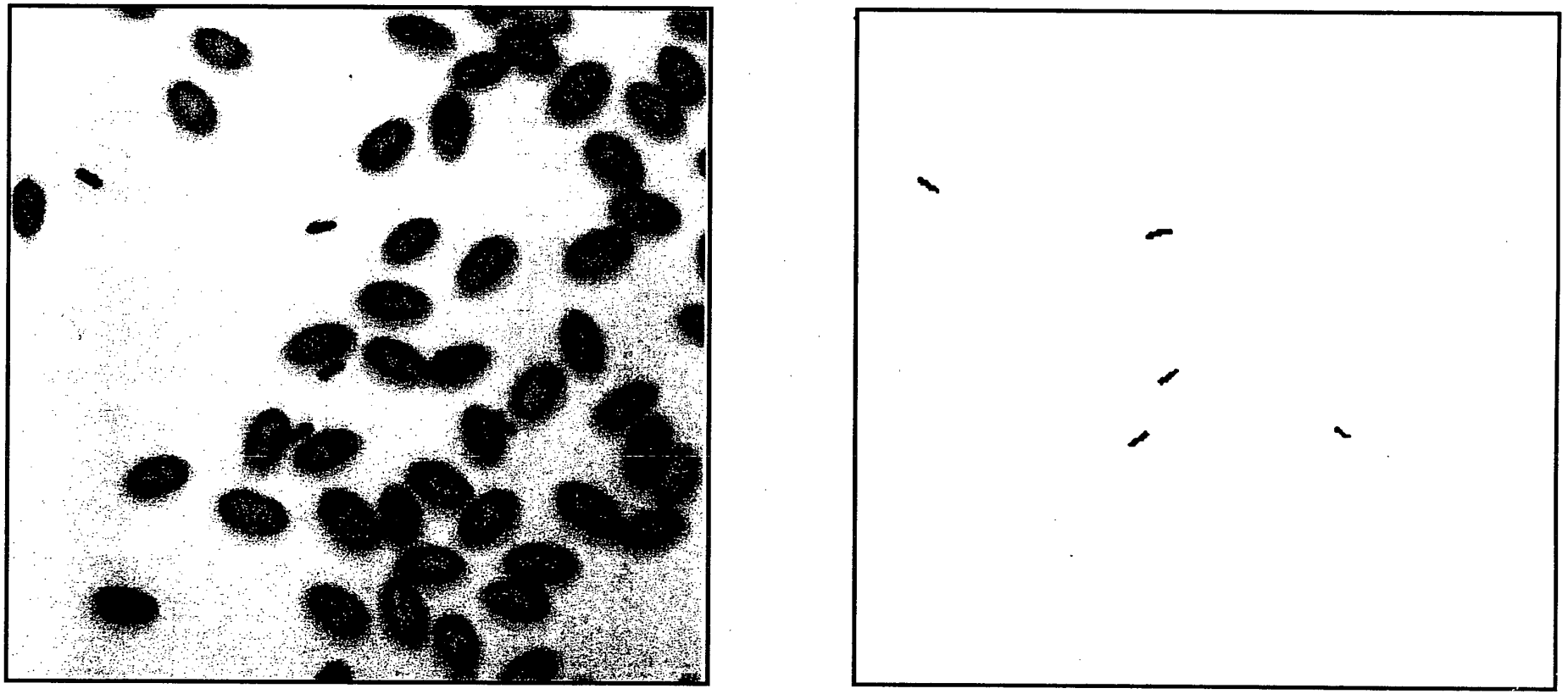

Figure 5: Example image containing adult *O. surinamensis* analysed by linear feature detector

Figure 6: Example image containing larval *O. surinamensis* analysed by the path tracking method

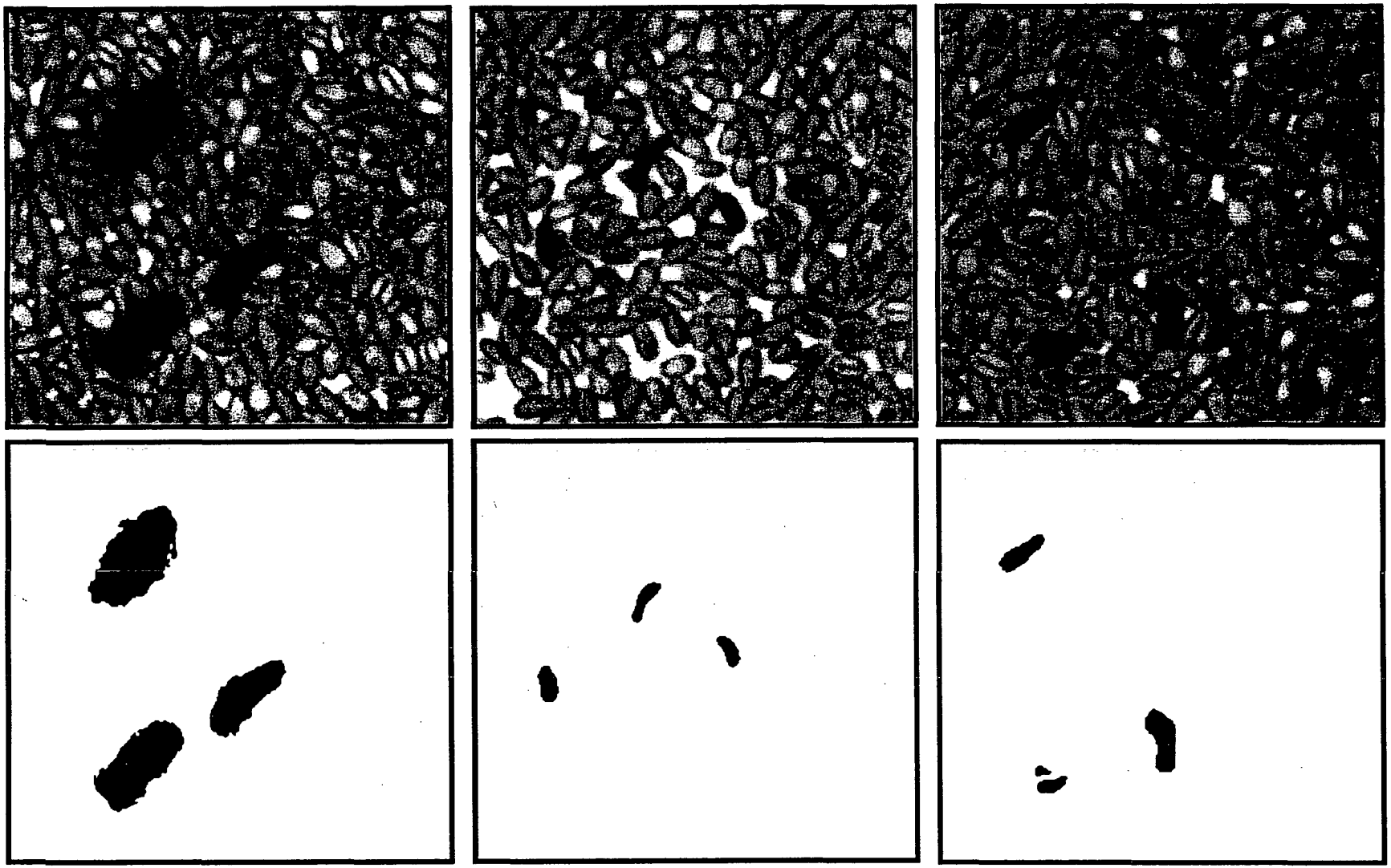

Figure 7: Example images containing (a) rodent droppings, (b) mouse droppings or (c) ergot analysed by the adaptive thresholding method

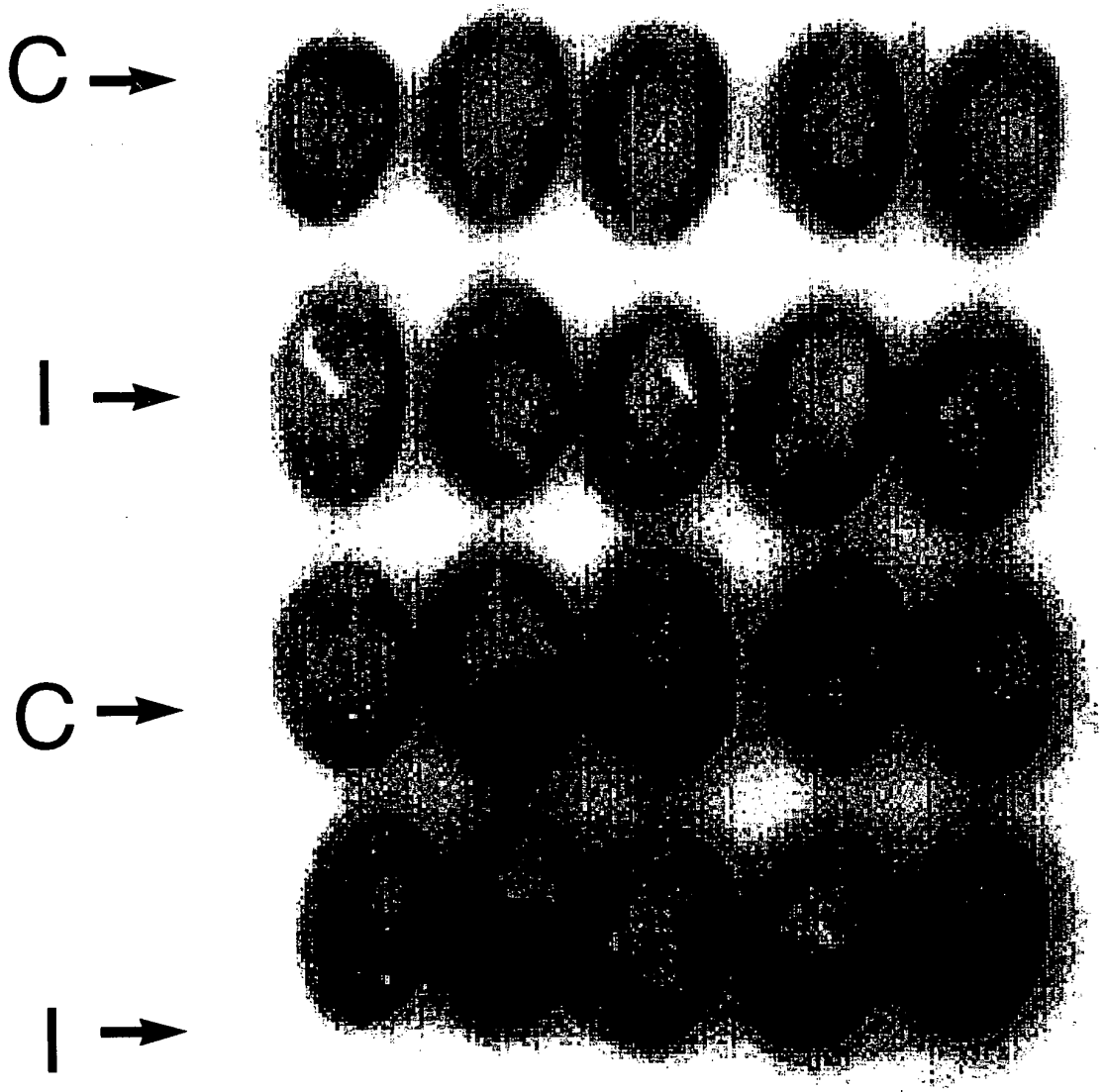

Figure 8: Experiment (n): image recorded in the visible region of wheat kernels internally infested with *S. granarius* larvae and control uninfested kernels

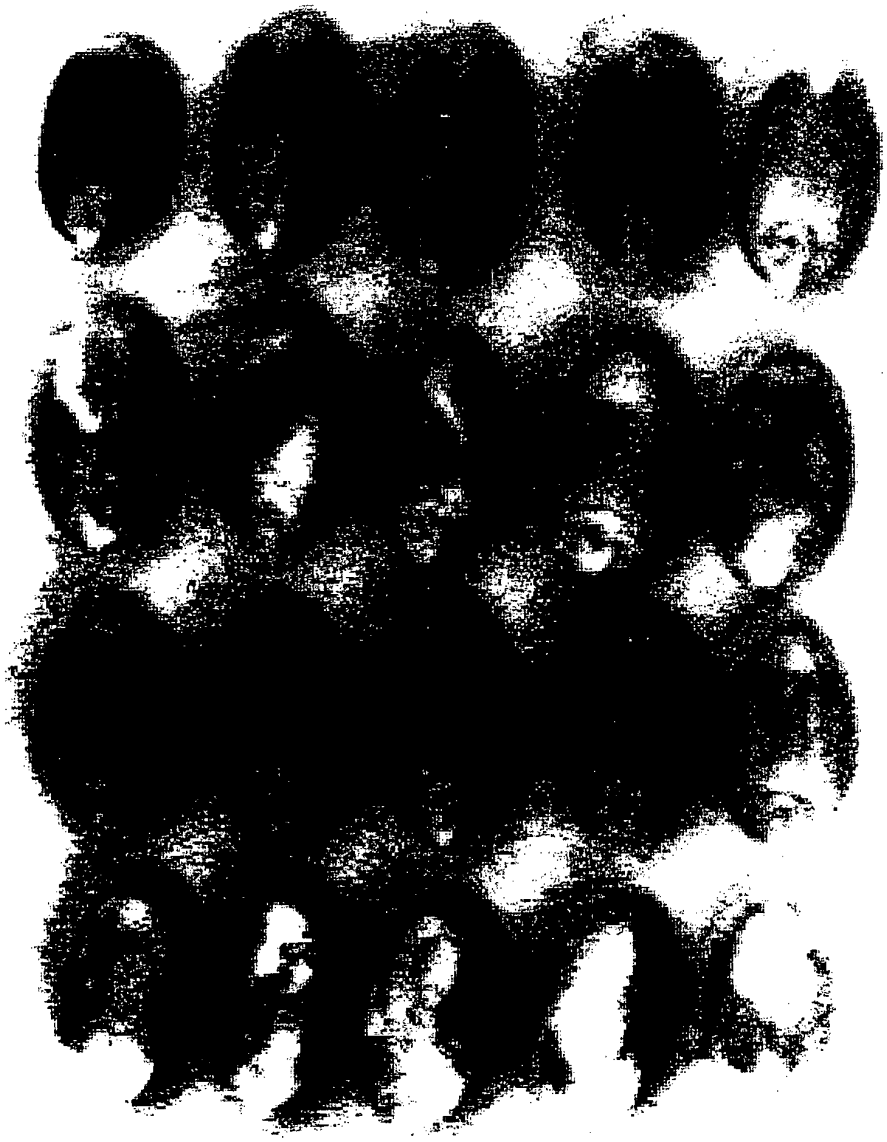

Figure 9: Experiment (n): image recorded at 1202 nm in the NIR region of wheat kernels internally infested with *S. granarius* larvae and control uninfested kernels

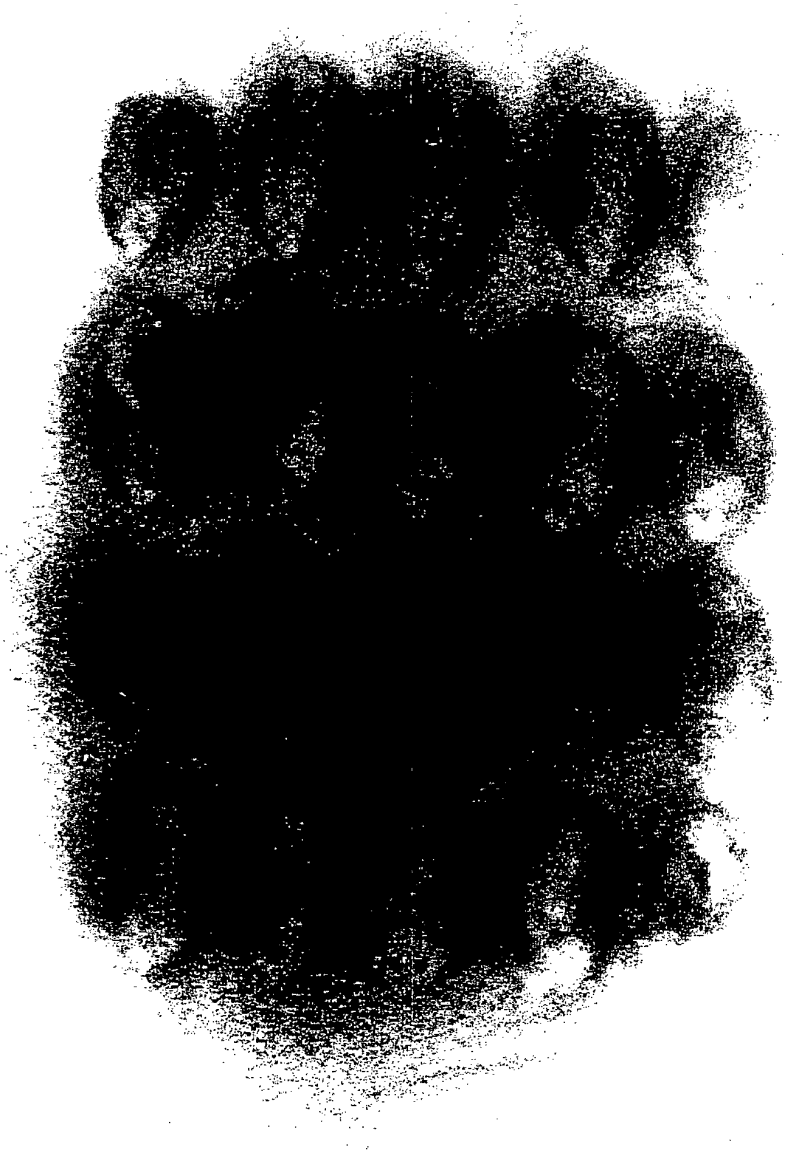

Figure 10: Experiment (n): image recorded 1300 nm in the NIR region of wheat kernels internally infested with *S. granarius* larvae and control uninfested kernels

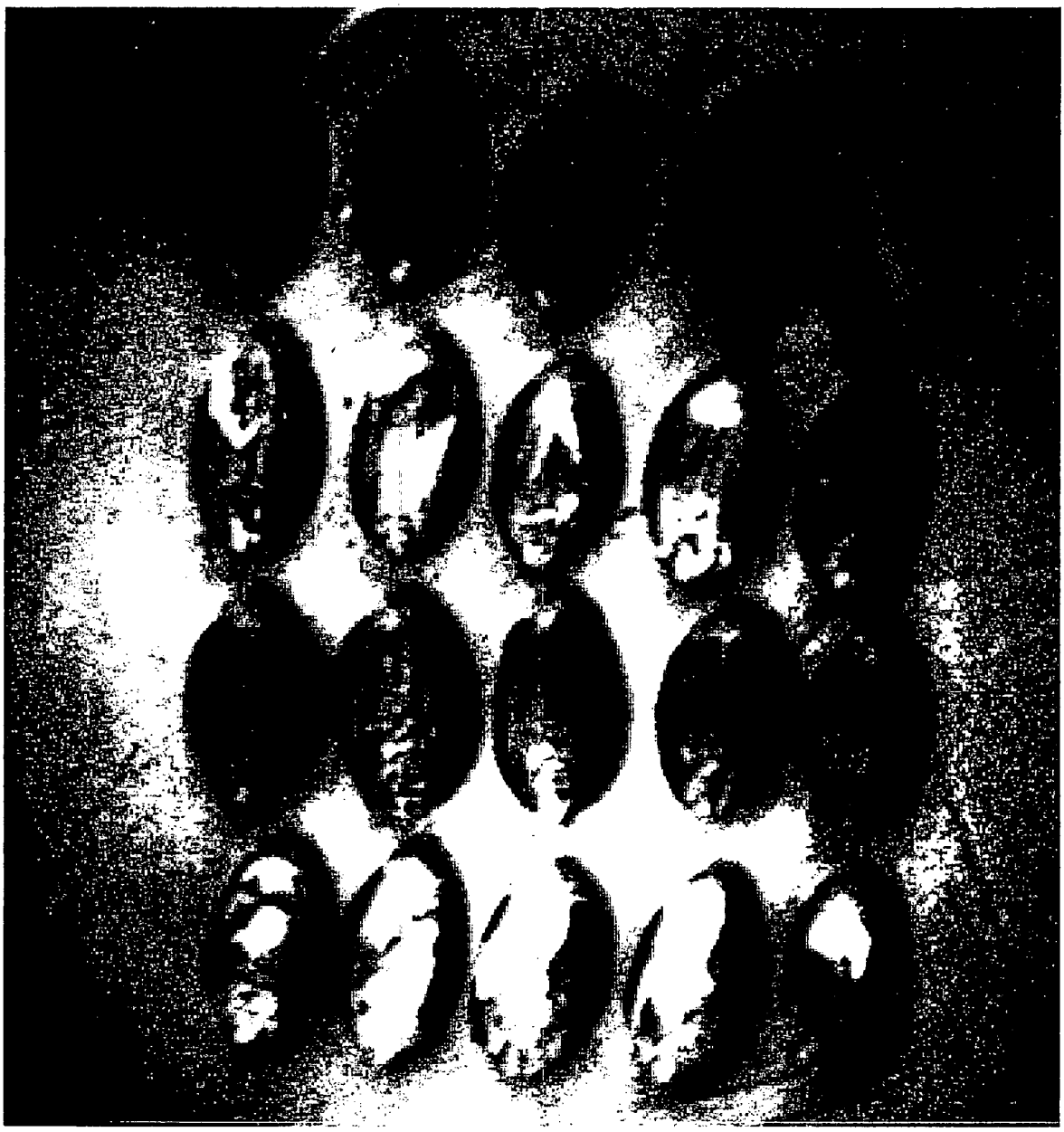

Figure 11: Experiment (n): difference image obtained by subtracting image at 1300 nm from that at 1202 nm

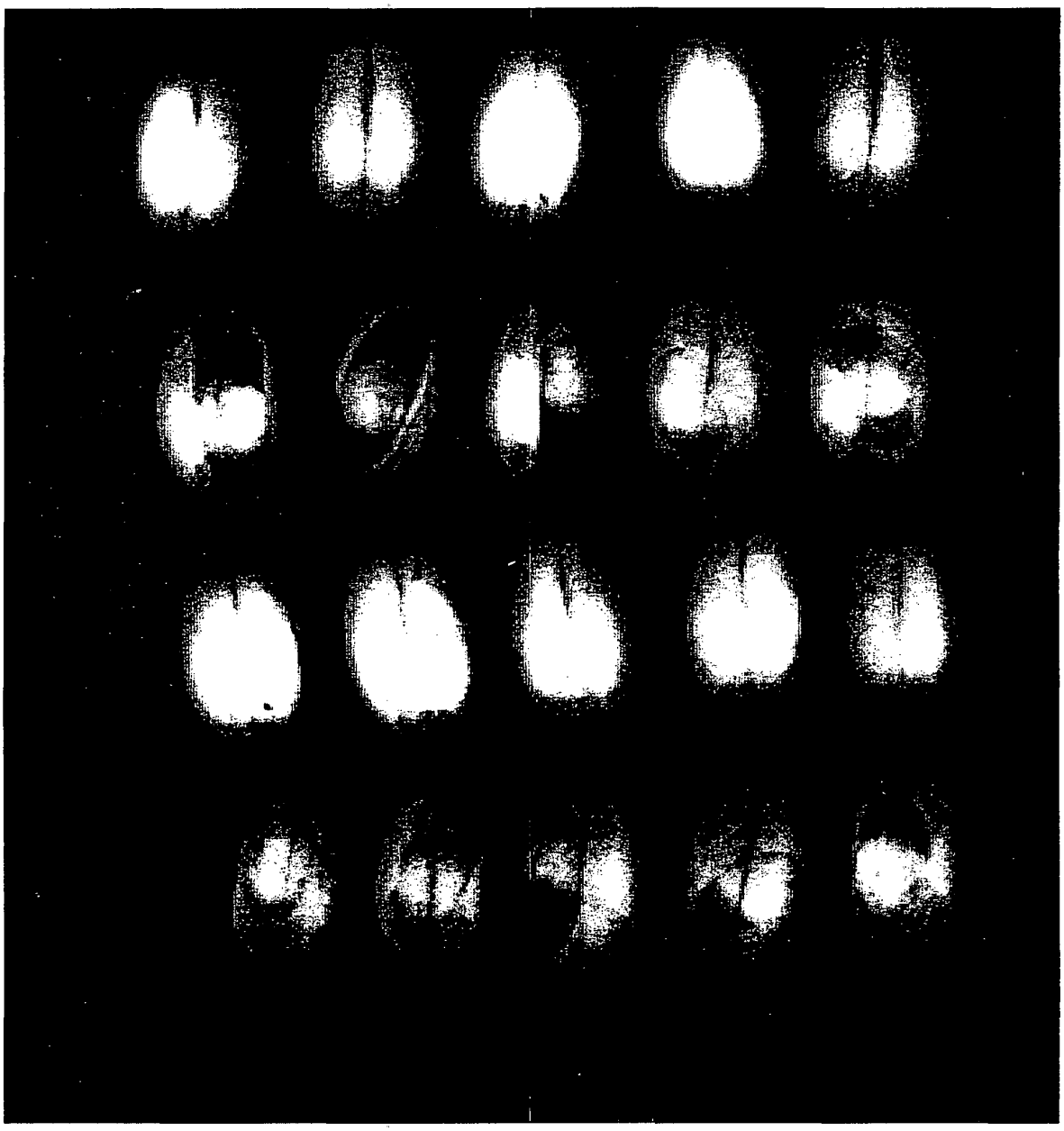

Figure 12: Experiment (n): image recorded in the X-ray region of wheat kernels internally infested with *S. granarius* larvae and control uninfested kernels

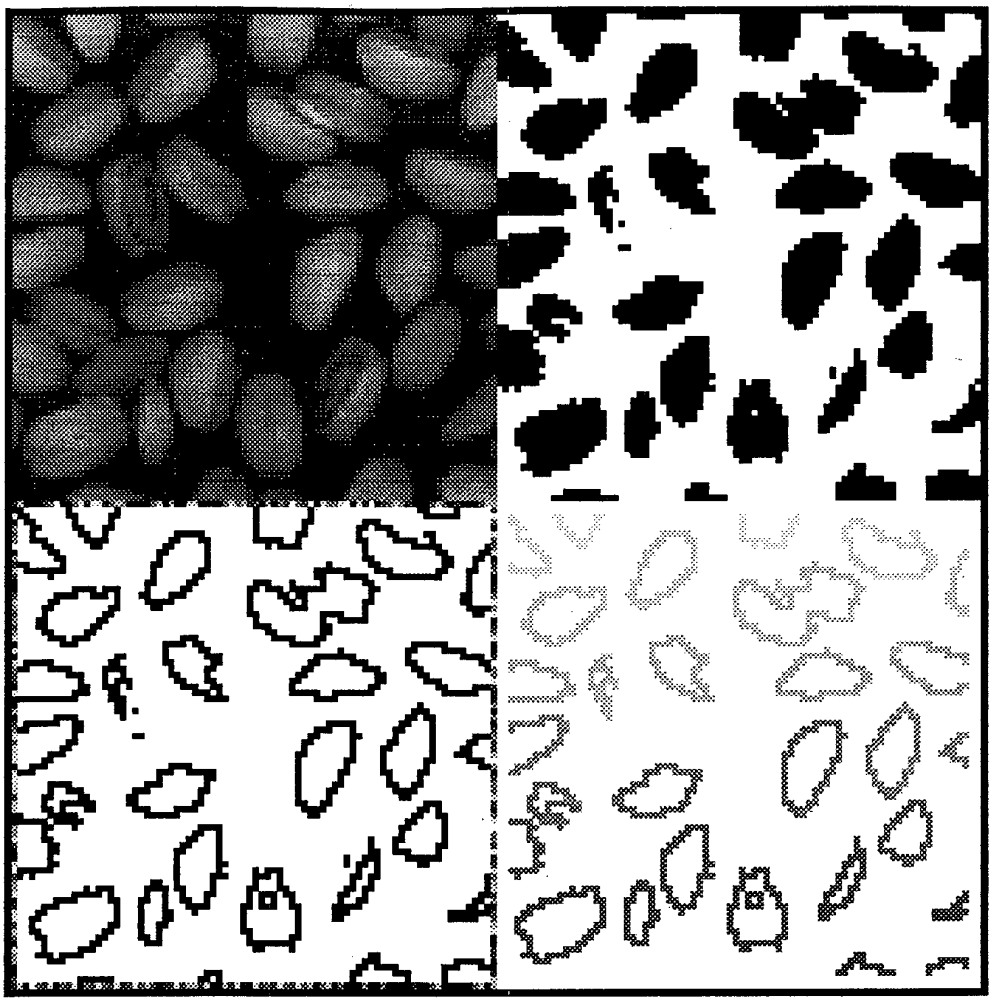

Figure 13: Experiment (p): image recorded in the X-ray region of wheat kernels internally infested with *S. granarius* adults analysed using the shape method. From top left clockwise: (a) original image including two infested kernels (b) thresholded objects (c) thresholded edges (d) labelled edges used in calculating object edge/object pixel ratio.

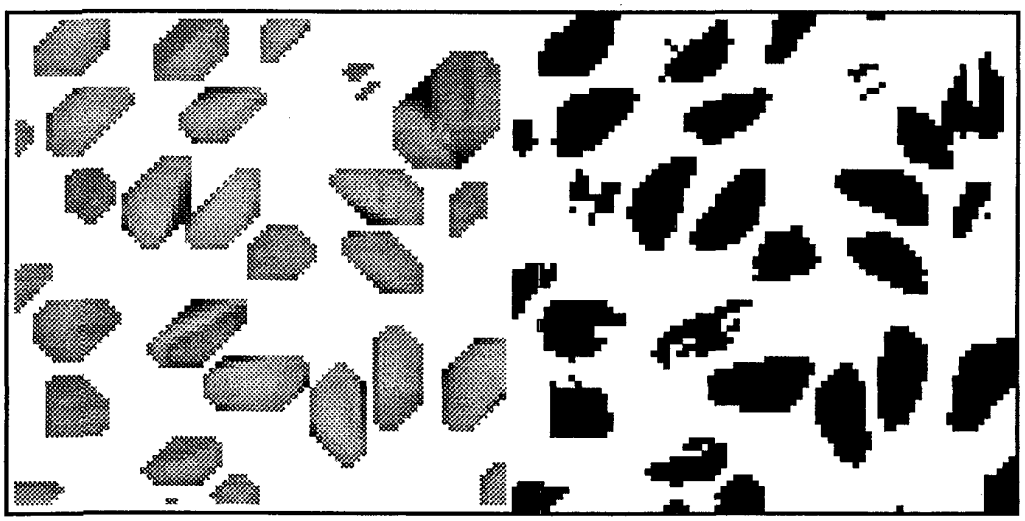

Figure 14: Experiment (p): image recorded in the X-ray region of wheat kernels internally infested with *S. granarius* adults analysed using the convex hull method. From left: (a) convex hull method to detect dark areas within kernels (insect feeding cavity) *versus* (b) thresholded image for input to shape method