Algorithmic recognition of biological objects

T. BERNIER and J.-A. LANDRY

Agricultural and Biosystems Engineering, McGill University, 21111 Lakeshore Road, Ste-Anne de Bellevue, QC, Canada H9X 3V9.
Received 12 March 1998; accepted 10 March 2000.

Bernier, T. and Landry, J.-A. 2000. **Algorithmic recognition of biological objects**. Can. Agric. Eng. **42**:101-109. An algorithmic method of object recognition to identify and count fungal spores in microscopic digital images is presented. The development of this process is a key element and cornerstone of a large-scale research program ultimately aimed at reducing fungicide application. The program, as a whole, is an attempt to build a machine based system in order to improve the ability of researchers to assess the population of pathogenic fungi within agricultural crops and thus more accurately target fungal pests. A three pass method was used: a preliminary pass in order to narrow the search space down to only the areas that contain spore-like darkening; a second pass that highlights both the center and the surrounding edge of the spore and produces a secondary image; and a third pass in which a template is matched to the secondary image. After the final pass, the list of positions and orientations of spores is reviewed and the conflicting and less likely positions are eliminated. The goal of the method is to accurately count the spores in the minimum amount of time. The resulting time is between 0 and 21 s of analysis on a 100 Mhz Pentium computer for a 64 by 64 pixel image. The algorithm, as implemented, demonstrated an accuracy of ± 5.3% on low quality images, which is less than the assumed error of humans performing the same task and is tolerant of partial occlusion. The system is loosely based on biological vision, is extremely versatile, and could be adapted for the recognition of virtually any object in a digitized image.

Le développement d'une nouvelle méthode algorithmique pour la reconnaissance de spores de champignons à partir d'images microscopiques est présenté. Le processus est une des composantes clés d'un programme de recherche global visant à identifier les spores de champignons accumulés sur des supports mécaniques. Cette information sera ensuite utilisée pour le développement et l'application de modèles de simulation épidémiologiques. La méthode présentée consiste en un procédé en trois étapes qui reconnaît avec succès les spores retrouvés dans n'importe qu'elle orientation, et est tolérant à l'occlusion partielle. L'algorithme, tel que développé, à démontré une précision de ± 5.3% lorsqu'utilisé avec des images de qualités médiocres, une performance bien supérieure à celle d'un humain dans des conditions similaires. La vitesse d'exécution s'est aussi avérée supérieure à celle d'un humain. La méthode développée présente un cadre descriptif qui, à travers les deux premières étapes, met en valeur certains critères distinctifs de l'image observée. Ces critères sont ensuite utilisés dans la troisième étape pour la reconnaissance finale des spores. La méthode est une adaptation libre du processus de vision biologique, est extrêmement versatile, et peut être adaptée pour la reconnaissance de presque tout objet présent dans une image digitalisée.

INTRODUCTION

This paper reports on a machine vision project that is the cornerstone of a large-scale research program ultimately aimed at reducing fungicide application. The program, as a whole, is an attempt to build a machine-based system to improve the assessment of pathogenic crop fungi populations and more accurately target fungal pests.

Pathogenic fungal spore density is considered to be a strong indicator of crop fungal infection (Vincelli and Lorbeer 1988). The determination of availability, dispersal rates, and dispersal patterns of spores is an important step toward the monitoring of plant diseases (Rotem 1988). Currently, the trapping and counting of airborne crop fungal spores is being used as an indicator of the presence and progression of disease. A problem with this approach, however, is that it can be very difficult to accurately assess pathogenic levels due both to variability in counts (Vincelli and Lorbeer 1988) and the difficulty to assess them while still allowing time for control measures to be taken.

The task of counting spores is time consuming and tedious. Due to the level of strain involved in this work, the workers can only perform for short periods with frequent cessations in order to rest their eyes and otherwise recuperate labour (Paul et al. 1993). Even at peak efficiency, manual counting of the fungal spore cells is performed too slowly. The data from one day of collection from a single sample location requires at least a full week to analyze. Thus, by the time the data is processed, the information is no longer relevant. In addition, it is generally recognized that accuracy in identification and counting decreases significantly after 3 to 4 h of labour (Paul et al. 1993). Automatically performing this task would reduce the amount of painstaking labour and also allow for the full analysis of the field samples within a useful time frame (optimally, in real-time). Thus, the goal of the overall research program is to develop a machine vision system that can count fungal spores in order to provide epidemiologists with an assessment of dispersal rates, distribution, and degree of colonization of fungus.

PROJECT OBJECTIVE and CONSTRAINTS

The objective of the project is to design an algorithmic process to identify and count spores in digital images.

The constraints are that this process:

1. be performed as fast or faster than by a human being,
2. have a precision of at least 80%,
3. be tolerant of partially occluded spores.

DEVELOPMENT OF THE METHOD and LITERATURE REVIEW

The problem involved developing a method to recognize irregular objects which may appear in any physical orientation and which may be obscured or occluded by other objects. The application of machine vision to agriculture is rapidly increasing and although there is a large amount of published

literature on the topic, there is none devoted specifically to the detection of fungal spores.

Machine vision in agriculture has traditionally been used in the grading of produce. One of the advantages of a mechanized process is that a machine can be absolutely consistent over long periods of time as opposed to the general inconsistency of and the effects of fatigue on human judgment criteria (Churchill et al. 1992).

Although grading applications of machine vision have met with reasonable success, many have two limitations in common: the position and orientation of the subject are controlled; and the subject is fully distinguished from both the background and other objects. Unfortunately, in the detection of spores, there can be no assumptions made about their position or orientation and occlusion is common.

Amongst the most common algorithmic approaches to the analysis of real images are variations of the Hough transform (Hough 1962), which has long been considered a robust and versatile technique for detecting analytically defined curves in natural objects. There have been many processes developed to detect and define circular or elliptical curves within images that are variations on the Hough transform (Raymond et al. 1992; Yuen et al. 1989). However, in these citations, a high degree of regularity in shape was assumed, thus an adaptation of this application to spore recognition is questionable.

A variety of attempts were made at image recognition using conventional methods and all met with similar failures. The failures seemed to be inherent to the processes themselves in that algorithmic detection processes are in fact, processes of detection. Thus, an attribute or parameter of a given object is sought within an image and any object that includes that given attribute is then considered to be the quarry. The frustration with the limitations of these methods in terms of natural objects in natural images led to the question: "Why is it so easy for humans to recognize objects and so very difficult for machines?" The processes of biological vision were then studied and its incredible generality inspired a much broader approach to the initial problem.

Vision in animals is extremely complex and not fully understood. We have a reasonable understanding of the processes involved in vision and the portions of our physiology and neurology that perform these functions. However, on the whole our understanding has many unanswered questions. It is in no way the intention here to explain biological vision but a brief introduction is necessary since a large portion of the recognition algorithm was loosely based on retinal functions.

Light entering the eyeball is focused by the cornea, pupil, and lens on the retina. The retina is a complex network of neurons that includes a mosaic of about 126.5 million photoreceptors. Each photoreceptor chemically transforms incident light into an electrical output signal. The response signal then passes through a network of cells and leaves the eyeball via the optic nerve, which is comprised of approximately one million ganglion cells. Early descriptions of the retina and its workings were made by Adrian and Matthews (1928). When more accurate means of measuring neural responses were available, the question of how several-hundred million receptor signals were mapped into only one million ganglion responses was addressed (Hartline 1940).

Experimentation demonstrated that there was a distinct spatiality in the grouping of the photoreceptor signals in the ganglion responses (Lettvin et al. 1959). It was also found that a single spot of light on the retina would elicit an action potential (response) from a particular ganglion cell. When that light was moved across the retina in a roughly circular pattern, the response of that ganglion cell remained the same (Kuffler 1953). In addition, if light fell on the area encompassed by the aforementioned circle, the response of the ganglion cell would cease even if light was still falling on the area that had previously excited a response.

From this information it was postulated that the signal of ganglion cells is actually an integration of the signals of many photoreceptors, which comprise a receptive field. There have since been many experiments looking into the grouping of receptive fields and it has been found that there are several varieties that respond to very specific stimuli and the concentric rings of mutually antagonistic responses is only a single example among many (Hammond 1973). In addition, this integration of signals is not restricted to the retina, in fact, the process happens throughout the visual pathway (Hubel and Weisel 1977).

It was decided to algorithmically emulate the retinal process of biological vision in order to create a far more generalized approach to recognition. The concept of the receptive field was the building block of the process and the overall approach was loosely based on the three cognitive processes humans undergo when searching for an object within an image: a swift scan; a closer look at the points of interest resulting from the preliminary scan; and finally a knowledge based decision as to whether or not the object in question is the sought object.

DESCRIPTION OF THE ALGORITHM

First pass: Initial glance

When searching for an object in an image, the tendency of most humans is to glance quickly through the image with a very coarse or crude aspect of the sought object as a search parameter, meaning that anything even remotely resembling a single or several given attributes of the sought object will require a closer look. However, until something similar is found, the search is relatively cosmetic, i.e. not every single location is compared with all the known attributes of the quarry but only the locations that contain a given aspect of the object.

The first pass of the algorithm behaves similarly. The receptive field (shown diagrammatically in Fig. 1) is passed over the image and at every sampled location, the average intensity of the pixels within the center and average of those in the surrounding ring are calculated, as depicted in the algorithm in Fig. 2. In locations where the center is significantly darker (a lower average intensity) it records the location, or in analogy to the biological equivalent, it responds.

After much adjusting of the sensitivity, the operation does tend to respond to some non-spore objects, however there are no omissions amongst the tested samples. Due to the irregularity in spores and the debris, it is essentially impossible to quickly respond exclusively to spores with no omissions, thus it was chosen to include some extraneous material rather

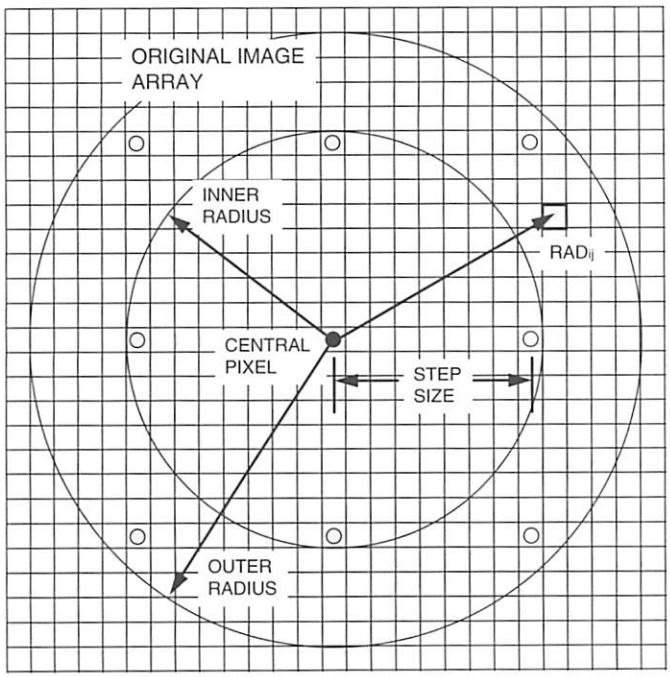


Fig. 1. Preliminary detection operator.

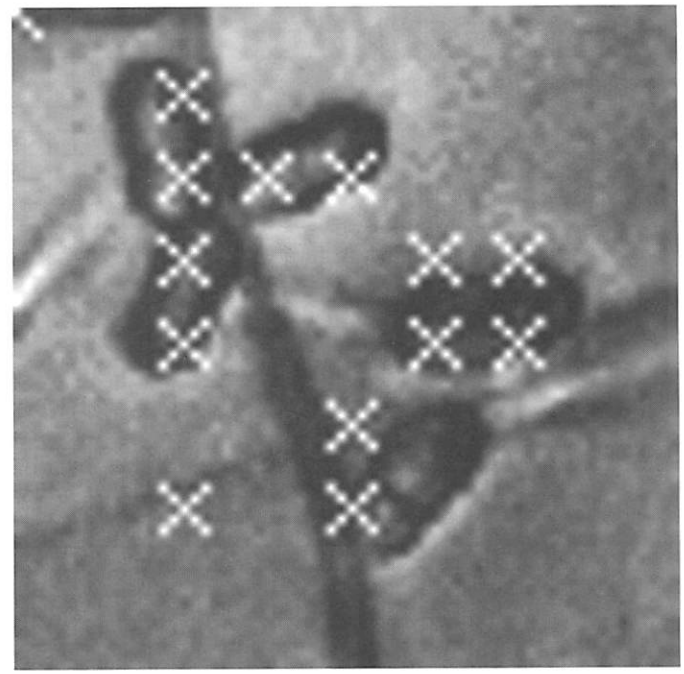


Fig. 3. Resulting responses of preliminary operator.

than possibly exclude spores. The response sensitivity is a matter of tuning with respect to the lighting conditions of the image; thus trials are required for best results. Optimal adjustment occurs when there is one response per spore.

With further experimentation and an effort at speed optimization, it was realized that it was unnecessary to apply the operator at every location within the image. It was found that by applying the operator only at intervals roughly equivalent to the smallest dimension of the spores, there were no extra omissions of spores and many of the extraneous responses of the operator could be eliminated. The resulting responses of the detection process are shown in Fig. 3. The X's in the figure are shown only to illustrate the points at which the operator responds and do not normally appear during execution.

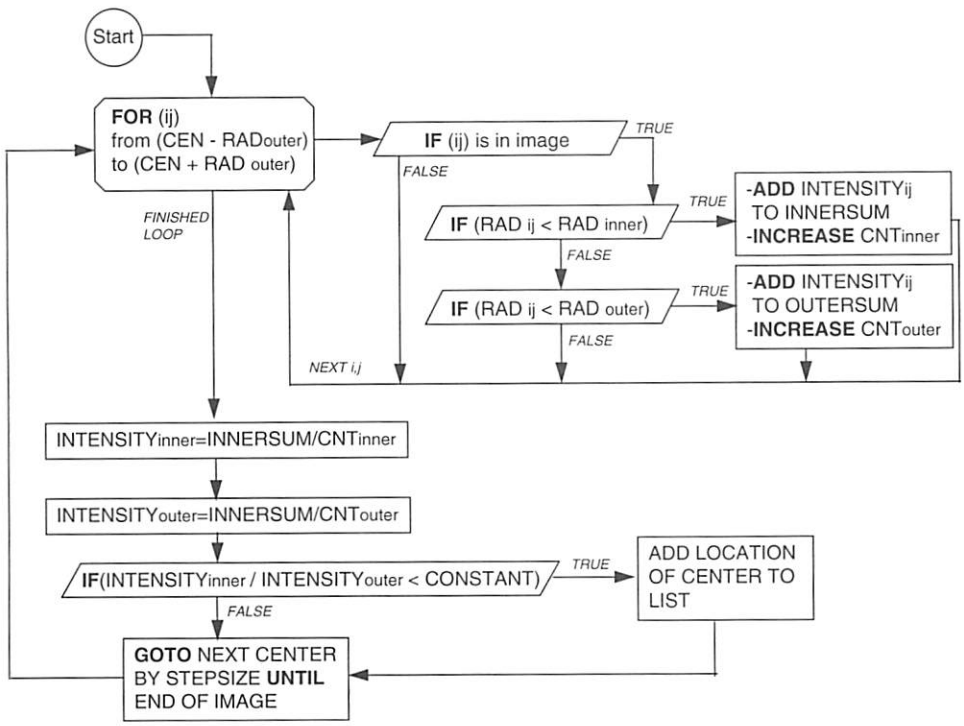


Fig. 2. Diagram of first pass algorithm.

Second pass: Local scrutiny

The second pass of the system is a considerably more object-specific process. While not a determination of location, it is an attempt to describe the image in terms of more object-specific qualities. This approach could be used on any two-dimensional object, however the tuning described here will be specific to the spores.

The receptive field of the second pass (Fig. 4) was designed to take into account some distinguishing features of the shading patterns of spores and only to search in the vicinities of the findings of the first pass using the algorithm described in Fig. 5. The field was designed to consider three separate zones: the center, the ring, and the periphery. Using these three zones of distinction and after much experimentation with various field sizes and response criteria, the most distinct outline of the spores was found (Fig. 6). The second pass lends regularity to highly irregular objects and is used as a means of better describing the image or

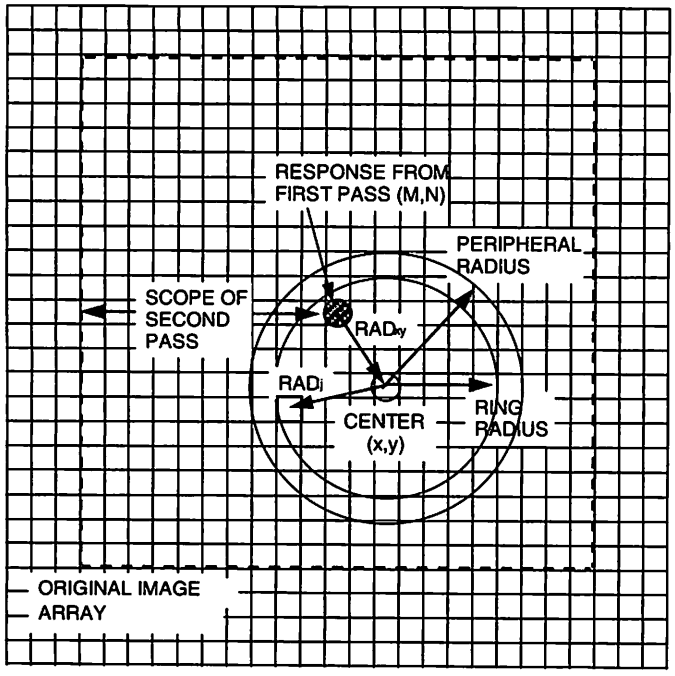


Fig. 4. Receptive field operator of second pass.

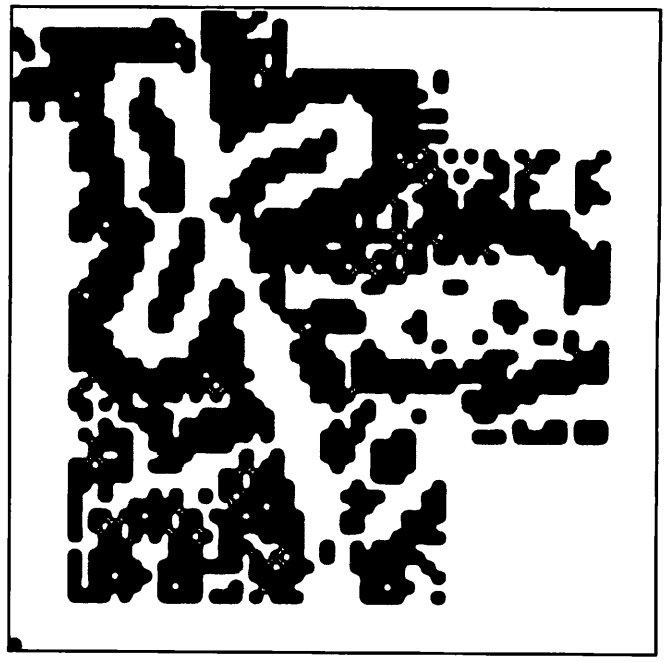


Fig. 6. Responses from selected tuning of second phase.

enhancing very general spore-like aspects of the image without being overly spore-specific.

Thus, the final result of the second pass is a strong response of the receptive field on the edges of the spores and a fairly dependable response on their centers. There is a very low response in between the edges and the centers, lending to a strong enhancement of the outline and center of the spores.

Third pass: Application of knowledge

While the first and second passes made an attempt at generality through an emulation of biological vision, the generality ends there. The third pass is very object specific. At this stage, what is known about spores is applied to their description resulting from the first two passes. The object-specific knowledge can be broken into two domains. The first is what the object looks like and the second is how the object can be arranged or more specifically, how it cannot be arranged. Its shape and its shading properties define what an object looks like. Thus at this stage, a shape and shading specific template can be applied to the image. Due to the radially asymmetric nature of the spores the template must be applied at every location and in every possible orientation. The template is applied to the second pass results and the orientation is coarsened to discrete increments.

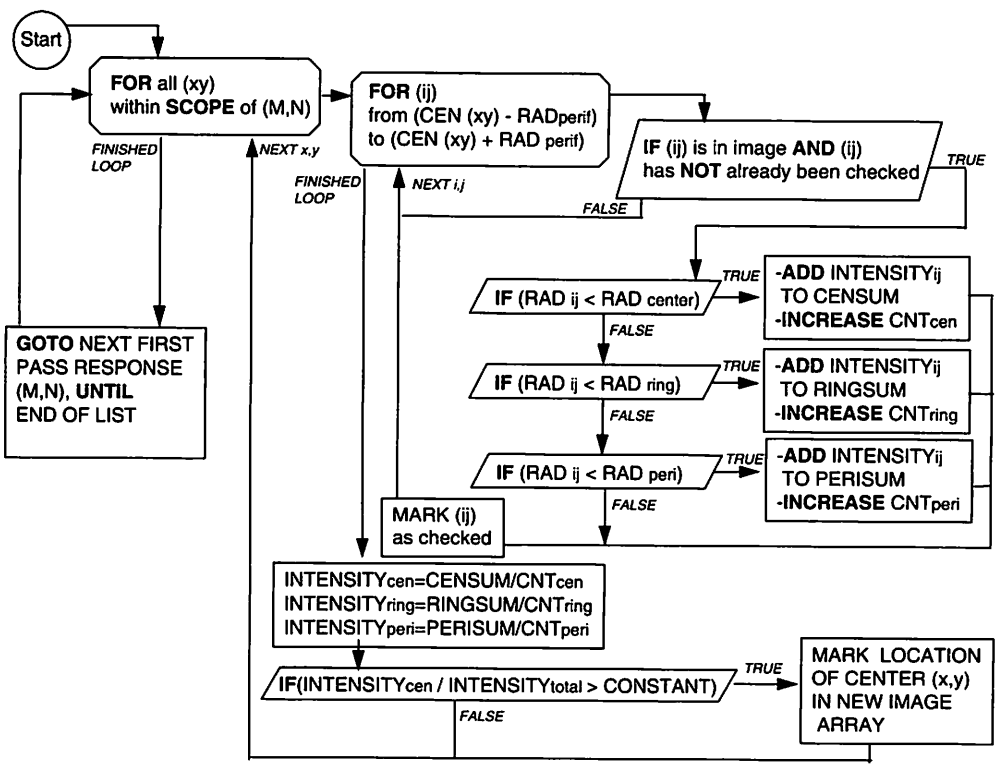


Fig. 5. Diagram of second pass algorithm.

However, a simple application of a template is inadequate. Although reduced, there is still a fair degree of irregularity in the representations of the objects and a simple match/mismatch criterion for a geometric shape would not work. The concept of a receptive field is therefore reintroduced and an

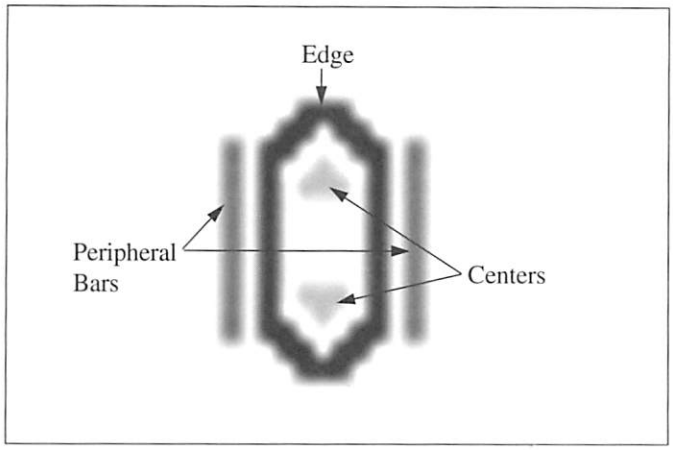


Fig. 7. Spore specific template.

asymmetric, shape and shading specific receptive field was constructed (Fig. 7).

The receptive field calculates, at every location and every orientation (Fig. 8) the matching of the peripheral bars to the typical highlighting of the contours of spores; the edge to the absence of a response; and the stronger aspects of the center response. For each of these aspects of matching, a score is calculated, using the algorithm shown in Fig. 9 and the positions and orientations of each strong match, above a selected threshold, are recorded.

The tendency of this matching process is to select more than a single orientation at any given location (Fig. 10). However, if the matching criteria were too strict, the poorer quality spores or partially occluded spores would be omitted, thus the selection of multiple orientations is left uninhibited.

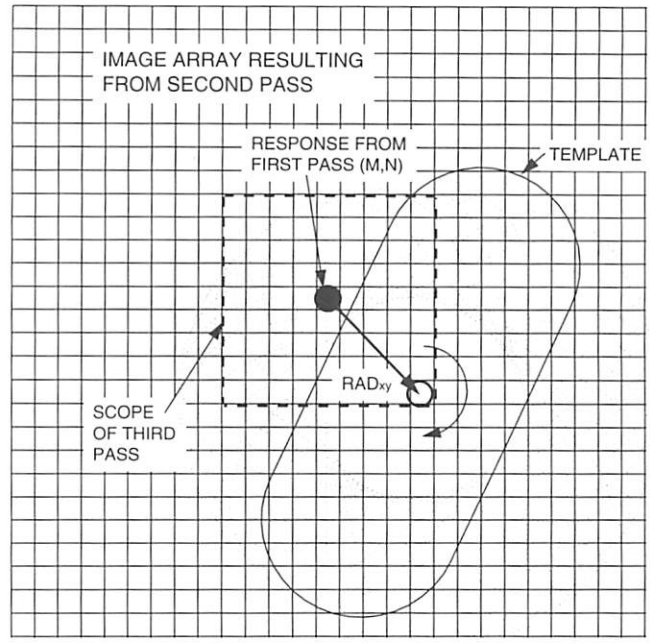


Fig. 8. Diagram of application of third pass.

This leaves the application of the second domain of knowledge, physical orientation. It must be recalled that although the process deals exclusively with a two-dimensional image, the image does represent a three dimensional object. Given that these particular objects are obloid and that they are collected by allowing airborne spores to fall onto the collection apparatus, the likelihood of having one spore lie directly on top of another (i.e. coincident centers) is very low.

Therefore, while occlusion must be tolerated, coincident centers must not be. With these precepts, the list of all recorded positions and orientations is re-examined and each recorded item is checked against its spatial relationship with all the others. For every item that conflicts with another, the matching scores are consulted and the strongest match is selected as the "correct" one. In addition, those matches that are weak and do not have any other matches in the vicinity that would possibly obscure them are discounted as non-spores. The final selections after the second domain of knowledge is applied are shown in Fig. 11.

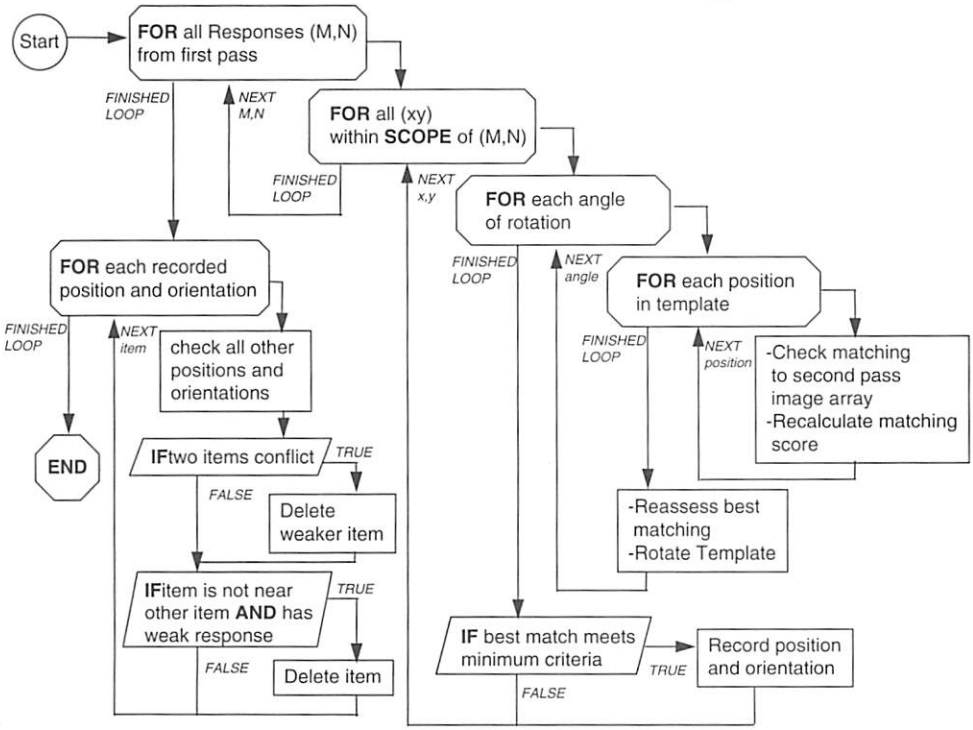


Fig. 9. Diagram of third pass algorithm.

MATERIALS and METHODS

The algorithm's code was written in C++ and was developed using a standard 16-bit development

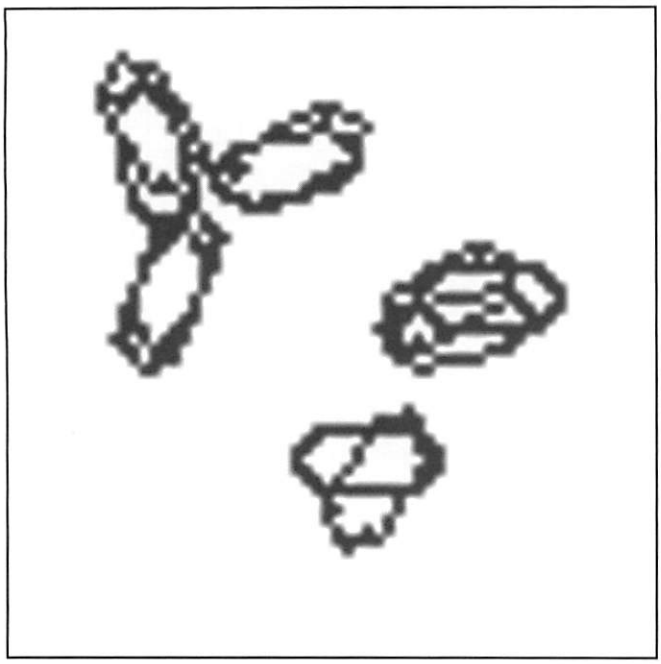


Fig. 10. Multiple matchings.

package. The language was chosen for its versatility and portability and the development platform for its availability at the time. Unfortunately, the choice of platforms led to some limitations. Due to the inherent 640 Kbyte scope of DOS and the 16-bit limit (64K index limit), the size of arrays used in the code had to remain relatively small. Since the images are most easily handled as arrays, the limitations were overcome by handling small portions (64×64 pixel) of the image at a time. In terms of development of the method, this has no effect. For an actual implementation or prototypical system, the code would require a 32-bit development process and a flat-memory platform such as Windows 95™ or OS/2™.

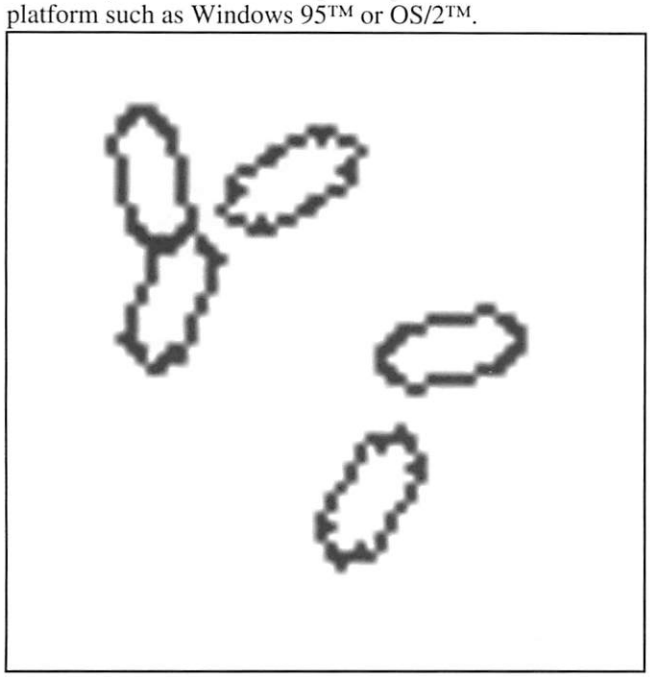


Fig. 11. Example of final result.

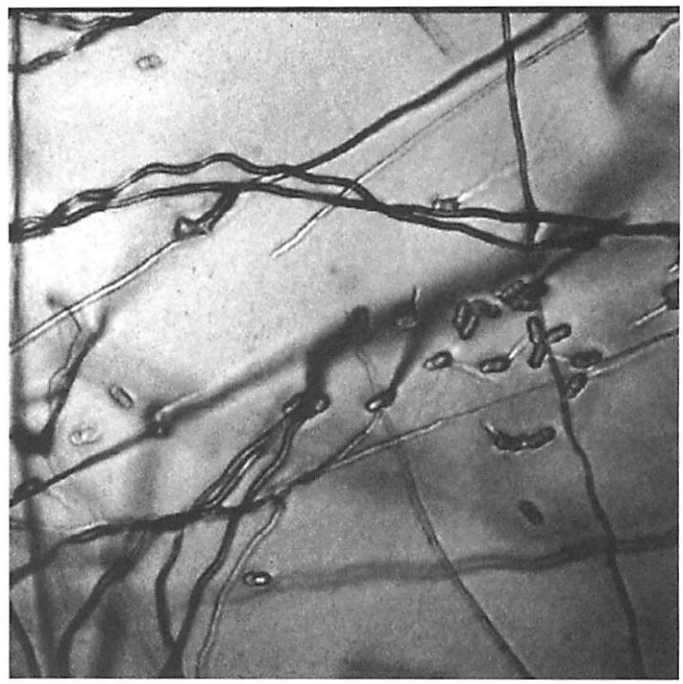
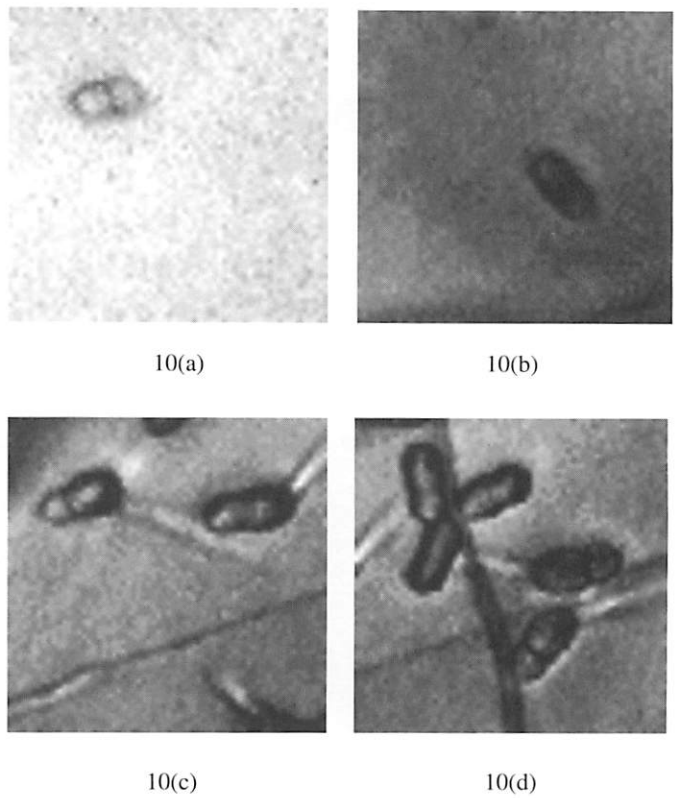


Fig. 12. Original image.

Due to the scope of this project, the image recognition system was developed before any actual *working* field samples were available. Thus a 400×400-pixel example image was provided in order to test the algorithm. The test image (Fig. 12), although displaying spores (the rice-like granules), differs



10(a) 10(b)

10(c) 10(d)

Fig. 13. Some selected 64×64 pixel test images.

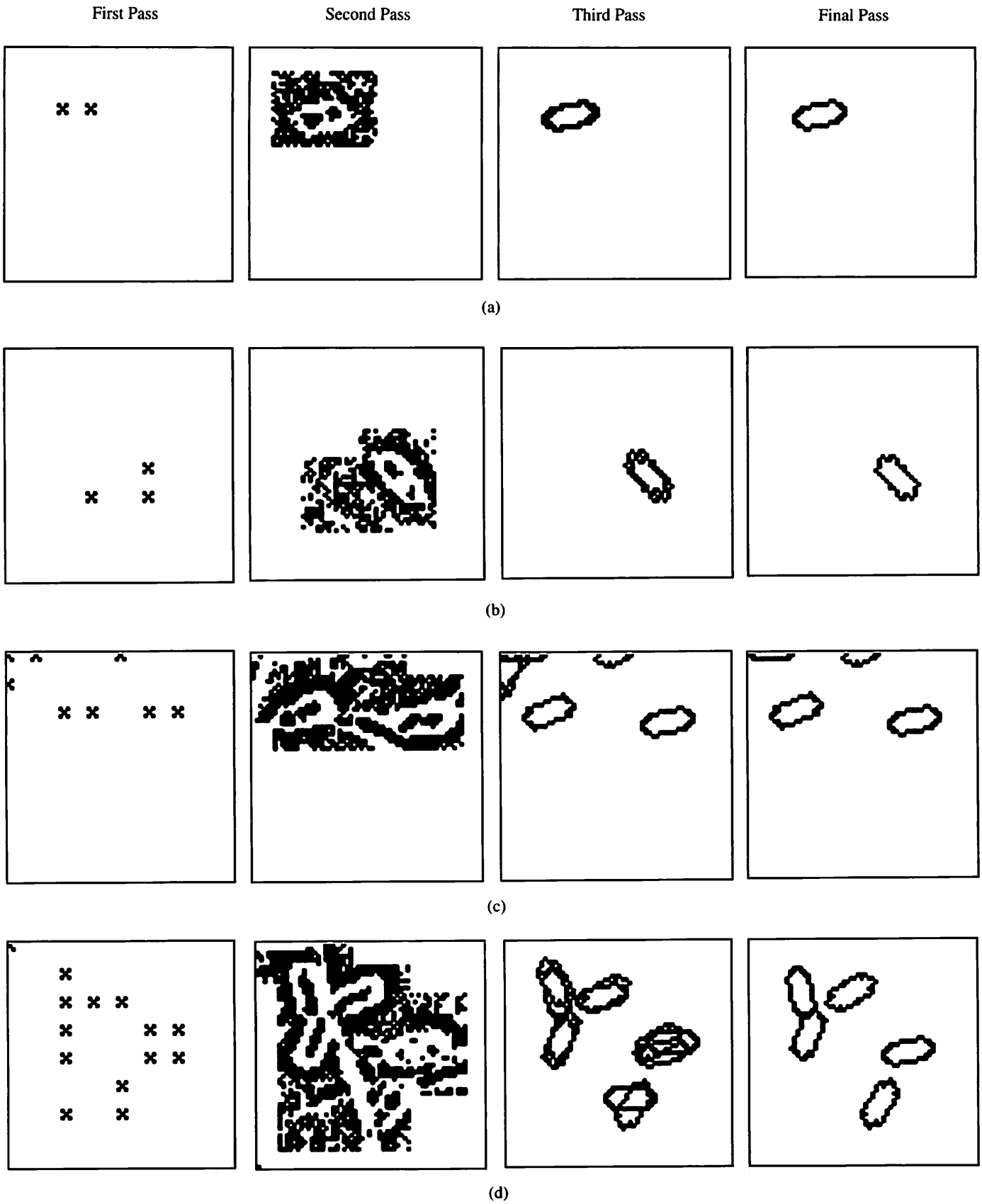


Fig. 14. Results of tests (on respective images in Fig 13).

significantly from an actual field sample. This image comes from a photograph taken through a microscope of a glass mounted spore sample, which was then digitally scanned. The sample, unfortunately, was fairly old when the photo was taken and as a result, there are cracks in the guar gum (long black lines). The concentration of spores is also considerably higher than would be in field samples and some of the spores in the image had germinated, which tends to obscure their shape. Furthermore, the lighting conditions across the image vary significantly. It was decided, however, that using this low quality image to test the method would lend to its tolerance of debris and occlusion and would yield a more robust algorithm.

The image selections shown in Fig. 13 were made on the basis that they represent a reasonable variety of conditions in terms of lighting and occlusion. While they are of relatively poor quality, the selections are not overly obscured by cracks or germinated spores and include concentrations of spores that could occur in field samples with the exception of Figure 13(d). This concentration of spores in Figure 13(d) is probably too high for such a small area but was included to thoroughly test the robustness of the process.

RESULTS and DISCUSSION

The developed process was applied to many test images and the results from the processing of the images in Fig. 13 are shown for each pass of the algorithm in Fig. 14.

There are two possible types of misidentification, false-positive (identification of a non-spore as a spore) and omission (not identifying a spore). In the above tests, there were no omissions, with the exception of the spores on the borders of the images in Figs. 13(c) and 13(d). The border spore omissions are trivial and could easily be avoided with field samples since there would be considerably less border per image (bigger images) and there could be a slight positional overlap as the image frame is moved over the sample to eliminate borders entirely. Counting a spore twice is also easily avoided since position and orientation uniquely identify each spore.

The false positive in the upper left corner of Fig. 13(c) is partially due to the border problem. When the receptive field straddles the border, only the pixels within the image frame are considered. While this generally leads to an omission, if the points that are considered confirm a spore and there is an inadequate contradiction, the false positive persists. Again, this is a border problem and can easily be solved.

Another false positive (not shown) did occur and is somewhat more significant. This is a misidentification due to the fact that a spore had germinated. While germination is an unlikely occurrence in field samples, this false-positive demonstrates that debris can fulfill the requirements of the operator. A solution to this problem would be to improve the operator and make it even more specific. However, increasing complexity in the operator would only further hinder the speed.

Overall the algorithm showed an accuracy of ± 5.3% on low quality images. In terms of speed, the estimated required time to analyze a sample of higher than likely spore density is approximately 68 consecutive days, while the required time for a sample of lower concentration by a human is approximately 58 working days or 80 consecutive days.

The three important differences with the images that will be obtained from field samples are: the quality, in terms of lighting and debris, will be improved; the resolution will be higher; and the handled images will be larger. All three of these factors will lend to a more accurate distinction of the spores. In addition, the algorithm will run under a 32-bit platform increasing its speed as well as its ability to handle very large arrays. Thus, the accuracy and speed will be improved.

CONCLUSIONS

The objective of this project was to develop an algorithmic process to identify fungal spores. Specifically, the process was to be fast, accurate, and tolerant of occlusion. All these requirements were met and although the test results were a little slower than expected, implementation on actual field samples will likely demonstrate adequate speed. The algorithm demonstrated a potential accuracy of ± 5.3% on very low quality images whereas the assumed error of humans is 20% even in ideal circumstances.

In terms of the machine vision project as a whole, even if this algorithm is not used in its implementation, it provides a great deal of insight into the problem. Since the algorithm is based on the retinal processes of biological vision, it provides a very modifiable means of describing any image, in terms of any features and an extremely adaptable template for the identification of any object in any orientation and the processes developed here could easily be applied to an artificial neural network based approach.

While the algorithm, as is, does not distinguish between physically similar species, it could easily be adapted to do so and currently provides important information in terms of position and orientation. Thus it can be seen as a valuable first step toward a system of species differentiation.

Future work on the project will include optimization of the code to improve performance, developing it on a 32-bit, flat memory platform, and extensive testing on field samples.

REFERENCES

Adrian, E.D. and R. Matthews. 1928. The action of light on the eye III: The interaction of retinal neurons. *Journal of Physiology (London)* 65(3):273-298.

Churchill, D.B., D.M. Bilsland and T.M. Cooper. 1993. Separation of mixed lots of tall fescue and ryegrass seed using machine vision. *Transactions of the ASAE* 36:1383-1386.

Hammond, P. 1973. Contrasts in spatial organization of receptive fields at geniculate and retinal levels: centre, surround and outer surround. *Journal of Physiology (London)* 228(1):115-137.

Hartline, H.K. 1940. The receptive fields of optic nerve fibers. *American Journal of Physiology* 130(4):690-699.

Hough, P.V.C. 1962. Method and means for recognizing complex pattern. US Patent no. 3 069 654, 1962.

Hubel, D. and T. Weisel. 1977. Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society (London)* Series B 198:1-59.

Kuffler, S.W. 1953. Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology* 16(1):37-68.

Lettvin, J., H. Maturana, W. McCulloch and W. Pitts. 1959. What the frog's eye tells the frog's brain. *Proceedings of the Institute of Radio Engineers* 47:1940-1951.

Paul, G.C., C.A. Kent and C.R. Thomas. 1993. Viability testing and characterization of germination of fungal spores by automatic image analysis. *Biotechnology & Bioengineering* 42:11-23.

Raymond, K.K.Y., K.S.T. Peter and N.K.L. Dennis 1992. A modification of Hough transform for circles and ellipses detection using a 2-dimensional array. *Pattern Recognition* 25:1007-1022.

Rotem J. 1988. Quantitative assessment of inoculum production, dispersal, deposition, survival and infectiousness in airborne diseases. In *Experimental Techniques in Plant Disease Epidemiology,* eds. J. Kranz and J. Rotem, 69-83. Berlin, Germany: Springer.

Vincelli, P.C. and J.W. Lorbeer. 1988. Forecasting spore episodes of Botrytis squamosa in commercial onion fields in New York. *Phytopathology* 78:966-70.

Yuen, H.K., J. Illingworth and J. Kittler. 1989. Detection partially occluded ellipses using the Hough transform. *Image and Vision Computing* 7:31-37.