Journal of Nematology 22(3):407-413. 1990.
© The Society of Nematologists 1990.

A Digital Image Analysis System for Comparing Groups of Small Nematodes

FRED G. PEET,[1] T. S. PANESAR,[2] T. S. SAHOTA,[1] AND JACK R. SUTHERLAND[1]

Abstract: A digital imaging system was developed for measuring various physical characteristics of individual nematodes and for comparing groups of nematodes. The equipment consists of a microscope, a video camera, a video digitizer, interactive displays, and a computer. Various physical and mathematical methods were incorporated, algorithms devised, and computer software written for image acquisition, editing, and analysis. To test the system, four populations of an isolate of the pinewood nematode, *Bursaphelenchus xylophilus,* subjected to 100% relative humidity at 22 C for 0, 12, 24, or 48 hours were compared. The results showed that the system can be used to measure physical parameters of individual nematodes and to differentiate groups of nematodes.
Key words: digital imaging, microscopy, nematode, pattern recognition.

A digital image is a matrix of numbers that represents the original object. For transmitted light, the individual numbers, called pixels (acronym for picture element), represent the amount of light that has passed through the object at each pixel location. Where the material in the optical path is less dense or lighter the number is relatively large, and where the material is denser or darker the number is correspondingly smaller. Because material surrounding the object is also digitized, it is necessary to isolate the object in the digital image, a step called editing. Once all the images have been edited, the resulting digital images of the objects can be analyzed. This involves determining the values of physical or mathematical variables, called features, which can be used to characterize different objects or groups, extracting the best features and performing multivariate statistical tests on them to determine if the various samples do indeed represent different groups.

Dusenbery (8,9) and Pline and Dusenbery (21) used a video camera and digitizer to track nematodes in response to different stimuli, but they did not use measurements to characterize the nematodes. Boag (4) used a digitizing tablet to make geometrical measurements on nematodes, but this did not involve digital imaging; rather, the operator traced the outlines of the nematodes with the tablet's cursor. Fortuner (11) developed a computer program to assist in the manual identification of groups of nematodes but did not use digital imaging. Panesar and Croll (15) and Yeates (23,24) used manual techniques to measure geometrical parameters of nematodes.

The objective of our study was to develop a system incorporating the techniques of digital imaging and pattern recognition to measure both geometrical characteristics of the nematodes and the probability distributions of stained material in the nematodes in order to obtain measures of the differences between populations. For this system it was possible to adapt and integrate some of the equipment and algorithms of digital imaging systems which were developed previously for studying insect cells and fungi (19,20) and which were based in part on the research of Bartels and Olson (3) and co-workers as cited therein.

MATERIALS AND METHODS

Data acquisition: The equipment consists of a Zeiss SMP-05 Scanning Microphotometer (Carl Zeiss, Oberkochen, West Germany), a Sony XC-77 black and white camera (Sony, Cypress, CA), an Electrohome EVM-15 black and white monitor (Electrohome, Kitchener, Ontario), a Gould FD5000 video digitizer and display (Gould,

Received for publication 10 October 1989.
[1] Research Scientists and [2]Postdoctoral Fellow, Pacific Forestry Centre, Forestry Canada, 506 West Burnside Road, Victoria, British Columbia, Canada V8Z 1M5.
We thank Ms. Diana Gundermann for digitizing and editing the specimens and Mr. G. A. Jensen and Dr. H. A. Moeck for assisting with microscope adaptations.

Fremont, CA), an Electrohome ECM-1301 monitor attached to the Gould display, and a DEC PDP 11/23+ microcomputer (Digital Equipment Corporation, Maynard, MA).

The microscope is fitted with an adapter for attaching the camera and a scanning stage with joystick and associated control electronics. The joystick allows the operator to move the slide for specimen selection. The signal from the camera is routed to the black and white monitor and then to the image digitizer. With this arrangement, that part of the object in the field of view is imaged on the monitor and allows the specimen to be viewed and focussed.

The Gould display is equipped with four 512 × 512 × 8 bit image memory banks, a feedback processor, an image digitizer, and a trackball. The feedback processor allows the video signal to be digitized 16 times, summed, and then divided by 16, all in a time of less than 1 second. This results in a digital image with less "noise" than an image formed from a single digitizing operation. (Noise is generated by random statistical fluctuations in the electrical signals and circuitry; it can increase or decrease the pixel values.) The output of the image digitizer is stored in one of the memory banks in the display. The digital image in the display memory is automatically directed, in the form of a video signal, to a monitor where it is displayed in a form resembling the original object.

The computer that controls and links this equipment together is equipped with 512 kilobytes of memory, two removable 10 megabyte disk drives, one fixed 26 megabyte disk drive, one removable 26 megabyte disk drive, a nine-track 1,600 bpi (byte per inch) tape drive, and floating point processor. The operating system is RSX11M-V4.1 (Digital Equipment Corporation, Maynard, MA).

The operator activates a terminal that controls the microscope and runs a menu-driven program. For a new population an index file is created; this file has the name of the population and ultimately contains a list of names which are assigned to nematodes that are digitized. Each day the camera output is digitized with and without light entering the camera. These two images are used for camera shading correction. For each nematode that is digitized, the population name, dark field name, bright field name, nematode name, time and date, and operator notes are recorded in the same computer file which ultimately contains the digital image.

The slide is placed on the microscope stage and the operator, with the aid of the joystick, moves the slide through the field of view. The slide is customarily moved in an S fashion until the whole slide has been viewed. That part of the slide in the microscope field of view is displayed on the black and white monitor; when a nematode is encountered, it is digitized by pressing the appropriate key on the terminal. The digital image is 512 × 480 pixels. Since the nematode occupies only a part of this area, the operator draws a line around the smaller region containing the nematode. This is done by viewing the digitized nematode on the monitor attached to the Gould display and using the Gould trackball to trace the line. The computer determines the minimum and maximum X and Y coordinates of this line, thereby defining a smaller rectangle within the larger digital image, and stores on disk this smaller digital image containing the nematode. The same procedures are repeated for each nematode within a population and for each population.

Data editing: The data acquisition step generates digital images that are rectangular in shape and enclose the nematode. Since the analysis is to be performed only on the nematode, it must be isolated within the digital image and the extraneous material outside the nematode removed. This is done by displaying the digital image on an Electrohome ECM-1301 monitor driven by a Datacube (Peabody, MA) QVG/AFA image digitizer and display. This equipment is attached to the PDP 11/23+ microcomputer. Although editing can be done on the Gould display, the Datacube is more amenable to this task. Conversely,

the original digitization could be done on the Datacube, but the Gould is more appropriate because it can sum and average digital images for noise reduction.

The operator runs the main nematode program which presents a menu. The image is displayed and at the same time the intensity level in the background is determined. This is used as a measure of the incident light intensity. The operator, through keyboard interaction, moves a cursor close to the nematode. The program searches in the neighborhood of the cursor for the edge of the nematode. Once the edge has been found, the program follows the outline of the nematode until the starting point is reached. This operation draws a line around the nematode and provides a measure of the perimeter (PERIM) of the nematode. The algorithm locates all points within the closed line and, when finished, has identified all pixels comprising the nematode. This operation provides the area (AREA) of the nematode. Then the light absorbancies of all pixels in the nematode are calculated. The absorbance is defined by:

$$\text{Abs} = -\log_{10}(I_t/I_i)$$

where I_t is the transmitted light intensity and I_i is the incident light intensity. The absorbance is proportional to the amount of absorbing material at the pixel location. From the individual pixel absorbances the total absorbance (TOTAB), average absorbance (AVGAB), standard deviation (STNDEV) of the absorbance, and maximum absorbance (MAXAB) are determined. If an absorbance has a value greater than 2.55, it is set to 2.55 which is well above the largest value found for stained nematode absorbance in practice. All the pixels outside the nematode are set to −1, and the edited image, in which pixels in the nematode correspond to absorbances, is stored on the computer disk. These operations are performed for each nematode in each population.

Data analysis—feature extraction: Features are mathematical or physical variables that potentially can be used to characterize a nematode population. Examples are perimeters, areas, total absorbances, average absorbances, standard deviations of the absorbances, and maximum absorbances that are determined during the editing step. Many other features can be computed. Presently, an additional 20 are determined; these are the probabilities of a pixel in the nematode having a particular absorbance range. This is done as follows. The absorbance scale is divided into 20 steps. The first 19 steps have a width of 0.04 and the last step has a width of 1.80. The scale starts at an absorbance of 0.0; therefore, the first 19 steps cover the absorbance range of 0.0 to 0.75; the last step covers the range 0.76 to 2.55. These step sizes and ranges were determined by examining the absorbances for several nematodes in two populations. The last step covers a wide range, since there were very few pixels above 0.75.

A histogram with 20 bins is determined for each nematode. Each bin contains a count of the number of pixels in the nematode which have an absorbance in the corresponding absorbance step. These counts are converted to probabilities by dividing by the total number of pixels in the nematode. Each probability is considered as a feature and is denoted by HISTnn, where nn takes on the value 01 through 20. For example, HIST01 is a feature that is the probability of a pixel in the nematode having an absorbance between 0.0 and 0.03 inclusive and HIST05 an absorbance between 0.16 and 0.19 inclusive. Including those features determined during the editing process, there are now a total of 26 features.

Data analysis—feature selection: Not all of the features are equal for characterizing the different populations; therefore, the features are necessarily ordered from most to least useful. This is accomplished by calculating a figure-of-merit (1,12,19,20) for each feature which is a measure of how useful the feature is for characterizing the populations. The figure-of-merit is the average of the ambiguity function (1,12,19), the measure of detectability derived from

TABLE 1. Two best features, corresponding figures-of-merit, Wilks's lambda, corresponding *F*-statistics with degrees-of-freedom, and Kruskal-Wallis statistic with degrees-of-freedom for each population of nematodes when compared against the control nematode population POP00.

Population	Best features	Figure-of-merit	Wilks's lambda	*F*-statistic	K-W
POP12	HIST01	0.43	0.84	11.9 (2,122)	28.4 (1)
	HIST07	0.38			
POP24	HIST12	0.59	0.44	73.9 (2,117)	68.5 (1)
	HIST04	0.58			
POP48	HIST12	0.76	0.24	170.4 (2,107)	79.7 (1)
	HIST03	0.73			

POP12, POP24, and POP48 are groups of nematodes exposed to 100% relative humidity for 12, 24, and 48 hours, respectively. Numbers in brackets are the degrees-of-freedom. The larger the figure-of-merit the better the feature is for differentiating between the indicated population and the control population. All the *F*-statistics and K-W statistics are significant ($P < 0.01$).

the receiver operating characteristic (1,19), and the correlations between features (1,20). It ranges from 0.0 to 1.0 with larger values representing better features. It is essentially a number indicating how far apart two distributions are when considered as functions of the feature under consideration. The features are sorted by their figure-of-merit value to obtain a list of the best 25.

Data analysis—statistical tests: The feature selection procedure provides a list of best features but does not indicate whether the populations are different. The latter question is determined by applying statistical tests to the data using the best features. The best features are considered as variables in a multivariate statistical analysis. Two tests are applied: 1) Wilks' lambda (2,22), and 2) a nonparametric test based on the Fisher discriminant function and the Kruskal-Wallis test (10,13,18). The first generates an *F*-statistic with appropriate degrees-of-freedom and can be checked for significance at desired confidence levels. If the *F*-statistic is significant, the groups of nematodes are not all the same. The tests assume multivariate normality and the equality of the covariance matrices. Programs are available to test these two assumptions for our populations (5,14). These assumptions are often not satisfied, but the results are reported because of the robustness of the tests (5). The second test does not rely on the above assumptions. It involves transforming multivariate data to univariate data using the Fisher discriminant function and then applying the Kruskal-Wallis nonparametric test (10,13,18) which generates a Chi-squared statistic. When working with two populations, the two groups of nematodes are said to be different if the resulting statistics are significant. When dealing with more than two populations, significant values of Wilks' lambda and the nonparametric test indicate only that not all the groups are the same. To resolve this situation, a generalization of the Fisher discriminant function is applied (7). This transforms the multivariate data into bivariate data in such a way that the distances between data points within the groups are minimized and the distances between groups are maximized. The means and covariance matrices for each group are calculated and confidence ellipses for each group are plotted based on the bivariate normal distribution.

Test nematodes: Nematodes from an isolate of the pinewood nematode, *Bursaphelenchus xylophilus* (Steiner and Buhrer) Nickle, from Clinton, British Columbia, Canada, were used to test the system. Details on the source and method of propagation of the nematode were given by Panesar and Sutherland (16). The nematodes were extracted in distilled water from 14-day-old cultures of the fungus *Schizophyllum commune* Fr. by the Baermann funnel technique and represented a mixture of all

TABLE 2. Four best features, corresponding figures-of-merit, Wilks' lambda, and corresponding *F*-statistic with degrees-of-freedom, and Kruskal-Wallis statistic with degrees-of-freedom when all four populations of nematodes are compared together.

Populations	Best features	Figure-of-merit	Wilks's lambda	*F*-statistic	K-W
POP00, POP12	HIST12	0.47	0.37	23.0 (12,614)	135.0 (3)
POP24, POP48	HIST11	0.46			
	HIST04	0.45			
	HIST13	0.43			

POP00 is the control population and POP12, POP24, and POP48 are the populations of nematodes exposed to 100% relative humidity for 12, 24, and 48 hours, respectively. Numbers in parentheses are degrees-of-freedom. The larger the figure-of-merit the better the corresponding feature is for differentiating the four populations. The *F*-statistic and K-W statistic are significant ($P < 0.01$).

developmental stages. Four samples from this set of nematodes were held separately in a minimal quantity of water in 10-ml beakers, ca. 4,000 nematodes in ca. 0.01-ml water. One sample was stained at 0 hours (control) in Oil Red O (17) and cleared in a water-glycerol mixture (6). The remaining three groups were placed at 100% relative humidity with no food source for 12, 24, or 48 hours at 22 C. At the end of their time periods, these nematodes were stained in the same way as the control group. The water was allowed to evaporate slowly at room temperature, leaving the nematodes in pure glycerine. Individual stained nematodes were selected randomly from each sample and mounted flat on clear slides in a minimal quantity of glycerine under Corning no. 1 coverslips. The absorbance of the stained nematodes was measured at 520 nm.

Fifty-eight stained nematodes from the control sample (POP00), 67 from the 12 hour sample (POP12), 62 from the 24 hour sample (POP24), and 52 from the 48 hour sample (POP48) were digitized and edited. Features were generated for each nematode, the best features were selected, and the above statistical tests were applied to the populations.

RESULTS AND DISCUSSION

The system was capable of easily measuring geometrical characteristics. It was found that the test populations did not differ with respect to the measured geometrical parameters. Nevertheless, the system can be used in those studies for which the geometry of the nematodes is important.

The probability distributions of stained materials in the nematodes were also easily measured. In comparisons of POP12, POP24, and POP48 against the control population, POP00, the best features were those derived from the probability distributions of the pixel absorbances in the nematodes. The two best features and their corresponding figures-of-merit, Wilks' lambda with corresponding *F*-statistic and degrees-of-freedom, and the Kruskal-Wallis statistic with its corresponding degrees-of-freedom for each comparison are listed in Table 1. All the statistics are significant ($P < 0.01$). The features are the relative frequencies of occurrence of pixel absorbances.

A comparison of POP00, POP12, POP24, and POP48 together was also performed. The four best features were HIST12, HIST11, HIST04 and HIST13. These and their corresponding figures-of-merit, Wilks' lambda with corresponding *F*-statistic and degrees-of-freedom, and the Kruskal-Wallis statistic with its corresponding degrees-of-freedom are listed in Table 2. All the statistics are significant ($P < 0.01$). Again, the best features were those derived from the probability distributions. A visual representation of these results is given in Figure 1. The multivariate data were transformed into bivariate data in such a way that the within-group distances were minimized and the between-group distances were maximized (7). The ellipses

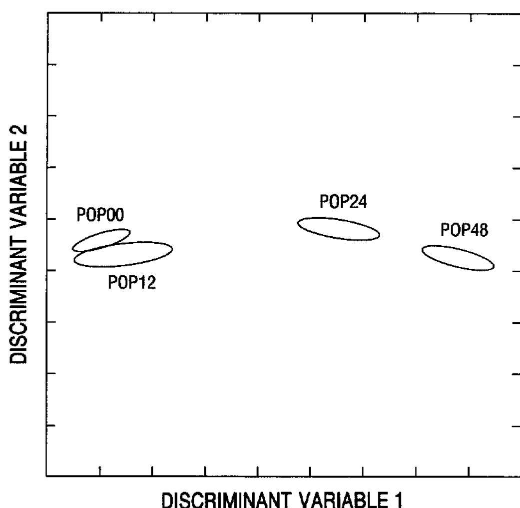

FIG. 1. Bivariate 95% confidence ellipses derived from the 4-variate HIST12, HIST11, HIST04, and HIST13 data for the four populations, POP00, POP12, POP24, and POP48, of *Bursaphelenchus xylophilus.*

were 95% confidence ellipses, illustrating that the four populations were different from each other. Discriminant variables 1 and 2 are the linear combinations of the best features listed in Table 2 that best illustrate the differences in the populations. In this experiment these differences arose from differences in the probability distributions of the absorbances of pixels in the nematodes.

Briefly, therefore, we have reported on a digital image analysis system that can perform geometric measurements on nematodes (4,15,23,24), obtain the relative quantities and the probability distributions of the stained material and make comparisons between populations based on these measurements.

LITERATURE CITED

1. Bartles, P. H. 1970. Numerical evaluation of cytologic data. 3. Selection of features for discrimination. Analytical and Quantitative Cytology 1:153–159.

2. Bartels, P. H. 1981. Numerical evaluation of cytologic data. 7. Multivariate significance tests. Analytical and Quantitative Cytology 3:1–8.

3. Bartels, P. H., and G. B. Olson. 1980. Computer analysis of lymphocyte images. Pp. 1–100 *in* N. Catsimpoolas, ed. Methods of cell separation vol. 3. New York: Plenum Press.

4. Boag, B. 1981. Measuring nematodes using a digitising tablet and microcomputer. Systematic Parasitology 2:145–147.

5. Box, G. E. P. 1949. A general distribution theory for a class of likelihood criteria. Biometrika 36: 317–346.

6. Croll, N. A. 1972. Energy utilization of infective *Ancylostoma tubaeforme* larvae. Parasitology 6:355–368.

7. Duda, R. O., and P. E. Hart. 1973. Pattern classification and scene analysis. New York: John Wiley and Sons. Pp. 114–228.

8. Dusenbery, D. B. 1985. Using a microcomputer and video camera to simultaneously track 25 animals. Computers in Biology and Medicine 15:169–175.

9. Dusenbery, D. B. 1985. Video camera-computer tracking of nematode *Caenorhabditis elegans* to record behavioral responses. Journal of Chemical Ecology 11:1239–1247.

10. Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems. Annals of Eugenics 7:179–188.

11. Fortuner, R. 1983. Computer assisted semiautomatic identification of *Helicotylenchus* species—the program Nemaid. California Plant Pest and Disease Report 2(2):45–48.

12. Genchi, H., and K. Mori. 1965. Evaluation and feature extraction on automatic pattern recognition system. Denkin-Tsushin Gakkai, Part 1 (in Japanese). Quoted in Tanaka, N., H. Ikeda, T. Ueno, M. Takahashi, Y. Imasato, S. Watanabe, and R. Kashida. 1977. Fundamental study of automatic cytoscreening for uterine cancer: I. Feature evaluation of the pattern recognition system. Acta Cytologica 21: 72–78.

13. Kruskal, W. H., and W. A. Wallis. 1952. Use of ranks in one criterion variance analysis. Journal of the American Statistical Association 47:583–621.

14. Mardia, K. V., J. T. Kent, and J. M. Bibby. 1979. Multivariate analysis. London: Academic Press. Pp. 148–149.

15. Panesar, T. S., and N. A. Croll. 1981. The hatching process in *Trichuris muris* (Nematoda: Trichuroidea). Canadian Journal of Zoology 59:621–628.

16. Panesar, T. S., and J. R. Sutherland. 1989. Pathogenicity of Canadian isolates of the *Bursaphelenchus xylophilus* (pinewood nematode) to provenances of *Pinus sylvestris* and *Pinus contorta* as grown in Finland: A greenhouse study. Scandinavian Journal of Forest Research 4:549–557.

17. Pearse, A. G. E. 1975. Histochemistry: Theoretical and applied, 3rd ed., vol. 1. Edinburgh, Great Britain: Churchill Livingstone/Longman Group Limited. Pp. 398–404, 418.

18. Peet, F. G., and T. S. Sahota. 1983. A nonparametric approach to the discrimination of two cell populations. Analytical and Quantitative Cytology 5: 201–203.

19. Peet, F. G., and T. S. Sahota. 1984. A computer assisted cell identification system. Analytical and Quantitative Cytology 6:59–70.

20. Peet, F. G., and T. S. Sahota. 1987. A digital image processing system for entomological studies. Proceedings of the IEEE Pacific Rim Conference on

Communications, Computers, and Signal Processing. Victoria, British Columbia, 1987. Pp. 258–263.

21. Pline, M., and D. B. Dusenbery. 1987. Responses of plant-parasitic nematode *Meloidogyne incognita* to carbon dioxide determined by video camera-computer tracking. Journal of Chemical Ecology 13: 873–888.

22. Wilks, S. S. 1932. Certain generalizations in the analysis of variance. Biometrika 24:471–494.

23. Yeates, G. W. 1972. Nematoda of a Danish beech forest. 1. Methods and general analysis. Oikos 23:178–189.

24. Yeates, G. W. 1988. Contribution of size classes to biovolume, with special reference to nematodes. Soil Biology and Biochemistry 20:771–773.