MAN-MACHINE INTERACTION IN PATTERN RECOGNITION

by

T. Kasvand
Control Systems Laboratory
Division of Mechanical Engineering
National Research Council
Ottawa 7, Ontario.

INTRODUCTION

Pattern recognition is nowadays the name applied to methods of extracting information from pictorial material. A subclass of pattern recognition, the so-called character recognition, is the name of the game if the objects to be extracted from the picture are alphanumeric characters. Reasonably successful commercial devices exist for reading typed or machine-printed material. Less success has been achieved in reading hand-printed alphanumeric characters. For general pictures containing "anything", the extraction of the desired information is surprisingly difficult.

Pattern recognition is becoming an increasingly important field of research since the amount of pictorial material to be analysed is increasing rapidly. Satellites are providing photos of cloud cover and terrain, aerial photographs give much detailed information on ground conditions, bubble chambers produce millions of photographs of atomic events, x-rays, γ-rays, etc. supply pictures for medicine and technology, the light microscope and the scanning electron microscope provide data on microstructure of organisms and materials, etc. Ideally we would like to have all this data analysed and classified by computers. However, due to our presently limited understanding of the nature of the pattern recognition problem, we can at present only program the computer to carry out certain routine tasks. A feasible solution to many pattern recognition problems is to let the machine do all it can on a given picture and, when it gets into difficulties, allow it to ask for advice from the operator.

THE NATURE OF PATTERN RECOGNITION PROBLEMS

To the "unitiated" pattern recognition problems appear extremely simple. He refers to his own experience of being able to recognize objects in his environment by just "glancing" at them. He can read printed and handwritten matter of any size and style even if the printing is upside down or if he has never seen an exactly similar writing style before. On a crowded street or in a windy forest his attention is immediately drawn to any object that moves differently from the way it is "supposed" to move, since such an object may imply danger or something of interest.

The possessor of such an elaborate information processing system, which he employs without even being aware of its existence, has of course difficulties in believing that the problem is complex. The literature on pattern recognition is full of examples where some trivial algorithm was proposed as the solution to some picture processing problem. After the algorithm was realized as a piece of hardware or software, only then did it become apparent there was "more" to the problem. Interestingly enough, one may now be able to change the problem to suit the algorithm, or to change the algorithm to at least solve a part of the problem.

For commercial character reading machinery the problem of reading characters has been modified to suit the algorithm upon which the reading is based. Thus, if the characters satisfy constraints on size, style, tilt, position, clarity of print etc then these machines work.

In most pattern recognition problems one cannot change the problem to suit the algorithm. On a microscope slide one cannot ask the bacteria to line themselves up like soldiers so that their counting and measuring will be easier, nor can one demand from a satellite photograph to show only the objects one is interested in and leave out the rest. Thus, in most pattern recognition problems we have very little control over picture content and only limited control over picture quality. The quality control consists mostly of not even trying to analyse pictures which are of too poor quality. However, one should not expect that rigid standards of quality can always be maintained. A practically usable pattern recognition system has to be able to analyse pictures of less than ideal quality. A limited control over picture content is obtained only if it is known at the moment of photographing what information is to be extracted from the picture. One may now be able to emphasize the desired information (or objects) at the expense of the remainder of the data.

It is usually safer to assume that one has limited control over the size of the desired objects in the picture and no control over their position and orientation. We may be able to throw away pictures where the desired objects bunch together, overlap, are partially hidden by other objects, are of poor contrast relative to the background, etc. If such pictures cannot be rejected then, whenever the computer gets into trouble, operator assistance should be requested.

Consequently, a presently feasible computer system for analysing pictorial material should have well-designed and easily understandable on-line displays for allowing the computer to display its problem, what it has done, where the difficulty is, etc., and allow the man to reply in a "comfortable" fashion.

TYPICAL COMPUTER SYSTEM

A typical computer system for pattern recognition and man-machine interaction consists of a scanner, a suitable computer and a set of interactive displays. The system in the Control Systems Laboratory of the Division of Mechanical Engineering at the National Research Council is shown in Block Diagram 1.

The scanner is a device which allows the computer to read the illumination (or gray) level at any point in the picture. It is highly desirable to allow the computer to read this value when it is needed, thereby treating the picture as a read-only memory for the computer. A flying spot scanner under computer control is the presently preferred device for presenting pictorial data to the computer.

The computer should be carefully selected since its size, popularity and the complexity of its operating system are of no significance compared to its "real-time" capability. A process control type computer with adequate memory, preferably with magnetic tape or disc backup, is often adequate. The computer is to contain the programs and the necessary data, but it should not be used to store the picture. It is normally thought that after the picture is scanned and stored on a computer-accessible medium (tape, disc) the major part of the pattern recognition problem is completed. This is not true. A picture resolved into say a 1000 x 1000 elements (each element represents at least one byte) seldom even fits into the memory of a large business type computer. Often the analysis consists of "looking" at each of these elements in the context of its neighbours (by using "local operators") and storing the results in one or more arrays of the same size. The final result of the analysis may only be a yes-no answer, a count, a histogram or a tabulation of results.

By treating the picture as a read-only memory, storage problems are alleviated and processing speed is considerably increased. Now we cannot mark off areas in the picture, nor can we get exactly the same gray level value during repeated reading from the same point in the picture. These problems, however, are easily overcome in the context of a given problem environment.

The selection of the display and interaction devices (like push-buttons, joystick, etc) is, of course, dependent upon what the operator is supposed to see and what he is expected to do. Consequently, a meaningful selection of man-machine interaction device can only be made in terms of a given problem. The generalities in this case only refer to the psychology of vision and the characteristics of the human operator.

GOALS IN MAN-MACHINE INTERACTION

Recognition of patterns is simple for ourselves. To obtain comparable performance from a computer program may be very difficult or impossible in light of present day knowledge and equipment. However, the computer is an excellent tool for carrying out accurate and rapid measurement or pictures. The goal of man-machine interaction for pattern recognition thus consists of making the optimum use of man's abilities as a pattern recognizer and the computer's ability to calculate. Fortunately these abilities complement each other.

The characteristics of man, as well as the nature of the interaction problem have to be considered when designing interactive displays. In the context of pattern recognition, where the computer uses the picture as a read-only memory, there is a minimum set of display requirements.

(1) Scan the whole picture and display it on a large screen memory scope. Now the operator can see the picture and there is a one-to-one correspondence between the picture on the display and as it is "seen" by the computer.

(2) The operator can now recognize the desired information and assist the computer if it gets into trouble or indicate the areas to be processed. The interaction consists in moving a pointer, a line, a circle, a box etc to the desired area, by pushing a button saying "take it", "quit what you are doing" etc.

(3) The operator should be able to call up a magnified display of any area on another memory scope, and to perform the interaction on the magnified picture.

(4) It is nearly mandatory to display the results of an algorithm (say contour following, centering, object separation etc) as a picture which overlays the previous scan. The operator can now see immediately whether or not the computer succeeded in accomplishing its task. If the operation was successful, the operator may want to see a "clean" display of the results.

(5) In general, a program for a picture processing task may or may not be able to detect whether it is in trouble. Even if the program "thinks" that it is "doing well", it should display its results. The attentive operator can now follow the progress and trigger an interrupt for manual interaction if necessary. If the program discovers that it is in trouble, it should immediately alert the operator by ringing a bell and produce the necessary displays to indicate the nature of the trouble.

(6) Since a practical pattern recognition system has to handle a large number of pictures of varying quality, one should program the system in such a manner that it does not fail immediately when the picture is less than perfect. Rather, the programs should have several levels of complexity. If the simplest algorithm - which is usually the quickest - fails, then the machine should fall back onto more elaborate algorithms before giving up and calling the operator. In this case the machine "realizes" that it is in trouble, and should start producing displays for the operator, while it tries to extract itself from the difficulty. If everything fails, or if the operator interrupts, the nature of the difficulty is already obvious to the operator.

SUMMARY

We take our ability to recognize objects in our environment as something obvious and think no more about it. Actually a very complicated information processing system is working for us which involves all our experiences, but we are not even aware of it. However, as soon as we try to solve a picture processing or a pattern recognition problem on a computer, only then are we beginning to realize the complexity of these problems.

The computer is excellent for carrying out accurate measurements on data in the pictures, but it may not always find this data. When the computer has exhausted its repertoire of algorithms, it should ask for operator assistance. By displaying the picture which the machine is analyzing, and by superimposing the results of computations, the operator can easily judge what to do to rescue the computer from its difficulties.

The observations described here are based on experiences with both practical and theoretical pattern recognition problems studied at the Control Systems Laboratory.

In many large research laboratories and universities picture processing laboratories have been established. Most of these systems are intended for specific picture processing problems. Descriptions of them may be found in the literature on pattern recognition.

APPENDIX

In pattern recognition the type of interaction between man and machine depends on the nature of the pattern recognition problem. In the simplest case the operator is only presented with the results of scanning and computations. He is only expected to check whether the initially set program parameters still apply.

In the recognition of nerve fiber cross-sections (Reference 1) the required picture size, its resolution and the bias level between black and white are to be set. For correct parameters the display appears as in Figures 1 and 2. If the bias is too high, the nerve fiber (the "rings" in Figure 1) become too thin and start to break up. The results of computing local curvatures of the contours is shown in Figure 2. If the fibers break up, curvature computations will fail.

It is often very convenient to initialize a picture processing problem via displays. In the study of human locomotion, (Reference 2) a walking man is filmed with a high speed camera which moves parallel to him. The man walks in front of a coded background. His limbs are marked with dots. The first frame of the film is scanned and displayed. The operator is now asked to point out where the code is located and which dots are to be used, see Figure 3. The location of the code is given by defining the x,y co-ordinates of the four crosses enclosing a part of the code. The operator moves a pointer to each of these positions and pushes a button. The dots on the limbs, the positions of which are to be measured, are pointed out similarly. However, in this case the computer also needs to know the size of the area to be searched, see circles enclosing the dots.

Figure 4 shows a display over a few frames of film. Figure 5 shows a long term display of limb motion during forward walk. The operator can now follow the progress of the program and guide it back onto the "right track" if a dot is confused with some shadow or becomes hidden by a limb.

Many photographs exist which have been produced for manual analysis. Since we are extremely good at recognizing pictures, we tend to want grid lines etc to facilitate our measurements. These details help us but can be a great nuisance in automatic analysis. In these cases it is often simplest to let the operator point out where approximately the required data is located.

Figure 6 illustrates a nose cone in a wind tunnel. The position, pointing angle and nose angle are wanted. In this case the operator is asked to point out the x,y axes of the picture and two sides of the cone. What he has to do is written at the edge of the screen. The x,y axes are scanned across repeatedly, the best fit is found and they are displayed on the picture over the original scan. The cone sides are then computed using least squares fit and displayed. The computer carries out the necessary calculations, prints and displays the data and gives a clean display of the cone sides which are compensated for film misalignment (see Figure 7). This display is filmed.

In more ambitious pattern recognition projects there usually is a large number of computer programs. If too many programs clamour for operator attention, the operator tends to become confused, he makes the wrong decisions, pushes the wrong buttons, etc. Thus the operator should not be overloaded with displays and requests.

In the chromosome recognition project (Reference 3) the operator interaction aspect is studied separately from the automatic processing. Figure 8 illustrates a coarse scan of a chromosome spread, while operator interaction is carried on. The gray level histogram is shown in Figure 9. He points out a chromosome in Figure 8, pushes a button for magnified display, Figure 10, and is now able to carry out measurements on it, simply "box it in", follow its contour etc. Whenever he desires, he can see what he has done by having the chromosomes numbered (Figure 11), test karyotyped (Figure 12), and so on.

On-line displays should be produced also for record keeping purposes. In a complicated pattern recognition situation even the person who conceived the procedures tends to lose contact with the machine and what the algorithms are actually doing, without a proper pictorial record. Various printouts, comments, error messages etc are handy but very much inferior to a good pictorial record of program behaviour. Many such displays from an on-line picture language program are shown in Reference 4.

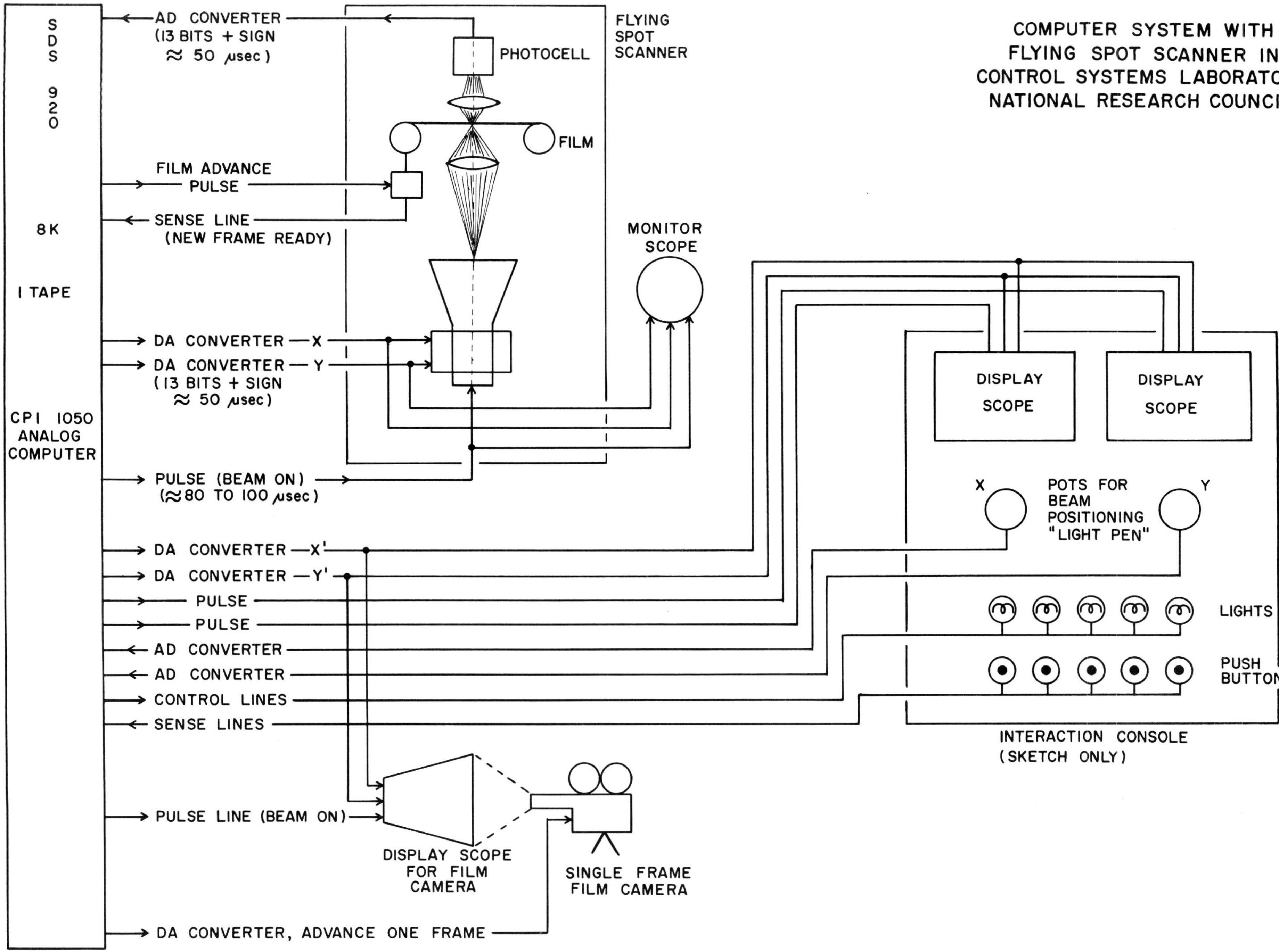
REFERENCES

1. Kasvand, T., Isnor, D., Romero-Sierra, C. — Man Versus Computer in the Measurement of Size Distribution of Nerve Fibers. National Research Council, Mechanical Engineering Report ML-5, January 1970.

2. Kasvand, T. — Pattern Recognition Applied to Measurement of Human Limb Positions during Movement. National Research Council, Division of Mechanical Engineering, Control Systems Laboratory Technical Report LTR-CS-41, November, 1970.

3. Hamill, P. — Experiments in Pattern Recognition - Current Work on Automatic Chromosome Analysis. National Research Council, Division of Mechanical Engineering, Control Systems Laboratory Technical Report LTR-CS-29, April, 1970.

4. Kasvand, T. — Experiments with an On-line Picture Language. Paper presented at International Conference on Frontiers of Pattern Recognition in Honolulu, Hawaii, January 18, 19, 20, 1971.



COMPUTER SYSTEM WITH FLYING SPOT SCANNER IN CONTROL SYSTEMS LABORATORY NATIONAL RESEARCH COUNCIL

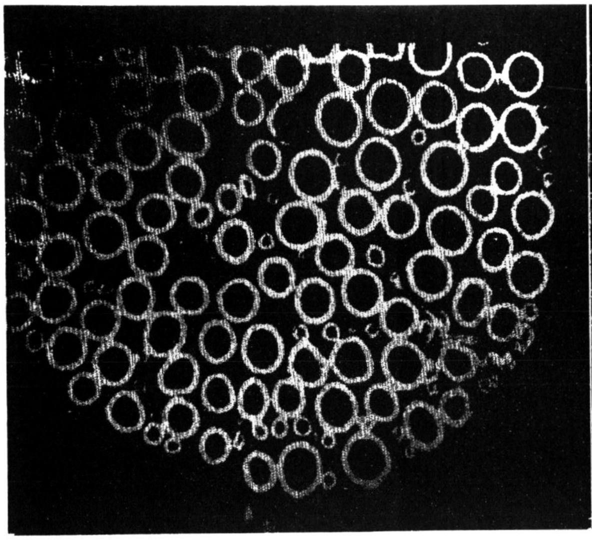


FIG. 1

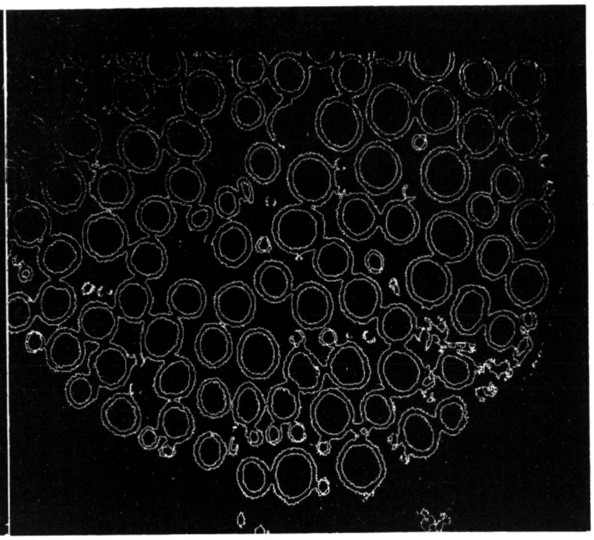


FIG. 2

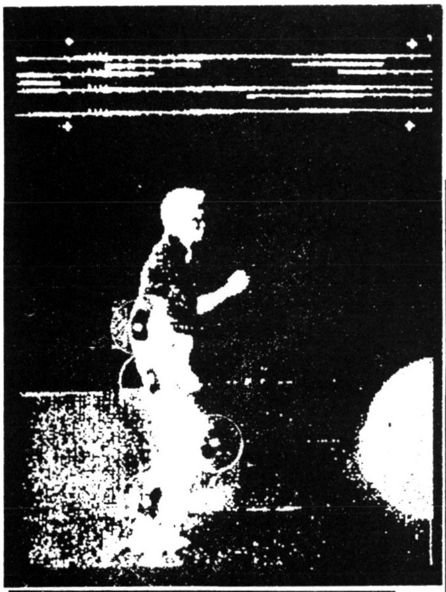


FIG. 3

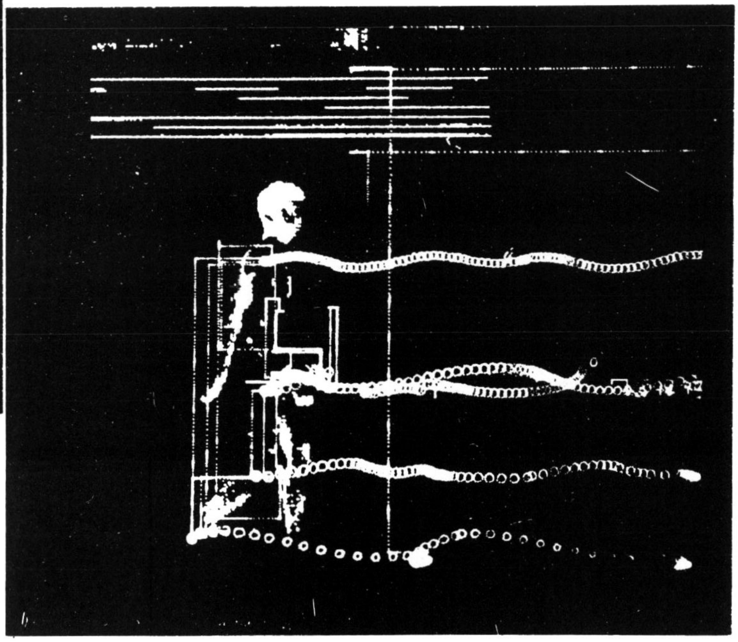


FIG. 5

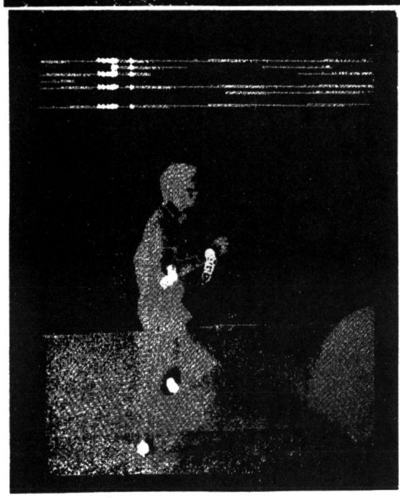


FIG. 4

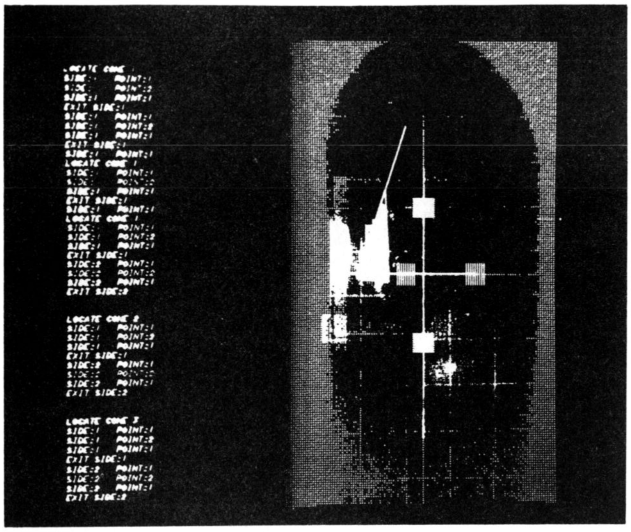


FIG. 6

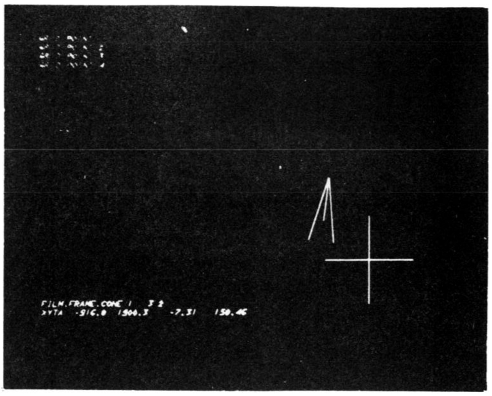


FIG. 7

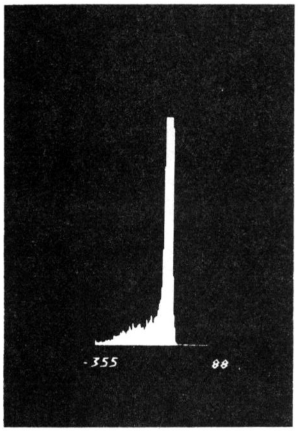


FIG. 9

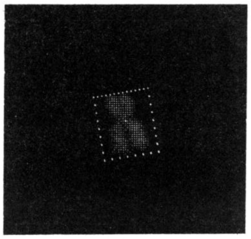


FIG. 10

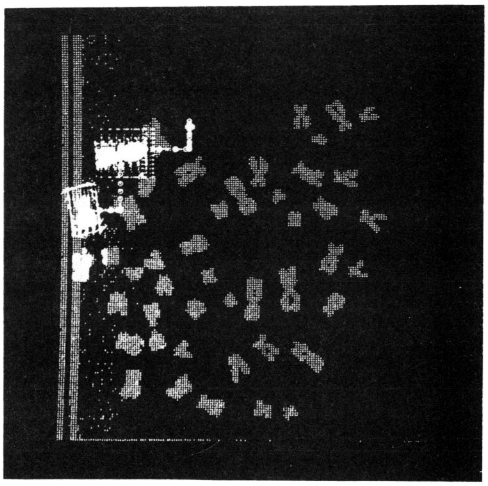


FIG. 8

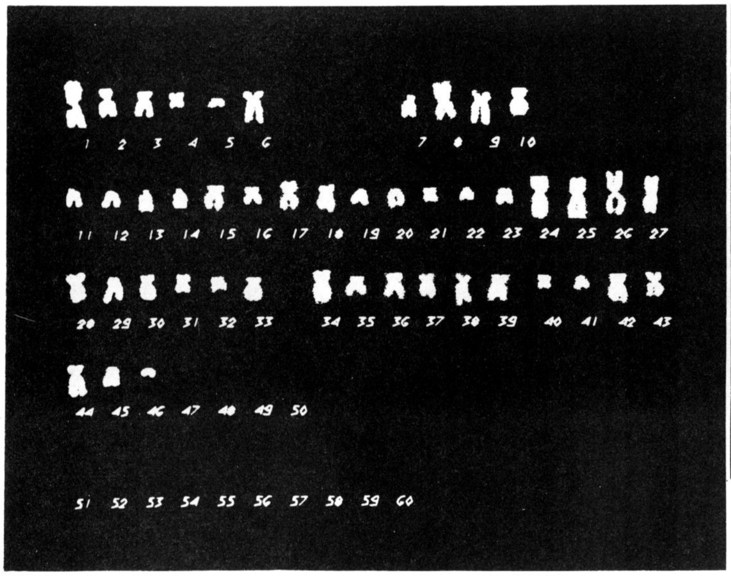


FIG. 12

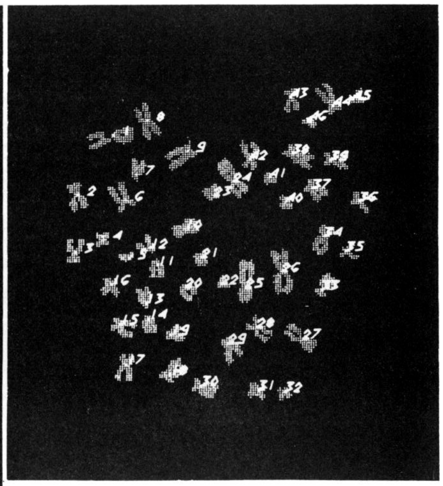


FIG. 11