MIT-T-74-001 c. 3

MIT
SEA
GRANT
PROGRAM

STUDY OF MEANS OF AUTOMATICALLY CLASSIFYING PLANKTON

by
Louis L. Sutro

CIRCULATING COPY
Sea Grant Depository

Massachusetts Institute of Technology

Cambridge, Massachusetts 02139

REPORT NO. MIT SG 74 - 11
JANUARY 31, 1974

STUDY OF MEANS OF AUTOMATICALLY CLASSIFYING PLANKTON

by
Louis L. Sutro

REPORT NO. MIT SG 74 -
INDEX NO. 74 - 311 - Vlk

CDSL REPORT R - 740

ACKNOWLEDGEMENT

The assistance of the following is gratefully acknowledged.

Dr. Ira Dyer, Director, MIT Sea Grant Program
Mr. Dean Horn, Executive Officer, MIT Sea Grant Program
Dr. Peter Wiebe, Woods Hole Oceanographic Institute
Dr. Paul E. Smith } National Marine Fisheries Service, South-
Dr. Elbert H. Allstrom } west Fisheries Center, LaJolla, Cal.
Dr. Max Petersen, Consultant in Optics
Prof. Hamish Munro, Dept. of Nutrition and Food Science, MIT
Prof. Ronald Hites, Dept. of Chemical Engineering, MIT
Prof. Whitman Richards, Dept. of Psychology, MIT

The research presented in this report was carried out by the Charles Stark Draper Laboratory, Inc., and sponsored through the MIT Sea Grant Program by NOAA Office of Sea Grant, U.S. Department of Commerce, under Grant number NG-43-72.

This report was prepared and published by the Charles Stark Draper Laboratory, Inc. as report R-740. Publication of this report does not constitute approval by the Department of Commerce of the findings or conclusions contained herein. It is published for the exchange and stimulation of ideas.

ABSTRACT

We have inquired into what appear to be three possible ways of classifying zooplankton chemically. One, classification by DNA, is too slow and expensive. A second method - grinding, pyrolysis, taking a gas chromatograph and a mass spectrogram - is not reliable. A third - analyzing the hydrocarbons in the stomach, gut and liver - has been applied only to large sea animal.

We have examined the way human sorters classify plankton to assist in developing a device that can perform in the same manner. Supporting this development are, first, a stereo-TV-camera-computer at the Draper Laboratory which can automatically form a three-dimensional model of a scene before it; second, designs being made at that laboratory of gimballed stereo TV cameras that can look around, as the human eyes and head do, with both high-resolution central views and lower resolution peripheral views; third, observation that human beings successively fixate, at the points where the binocular axes of the cameras converge, features on the surfaces of objects, and process both the features and the scan path between them, to recognize objects; fourth, solid-state sensor arrays to transduce the image; fifth, increasingly small, fast and inexpensive computer components.

Since the needs of industry for an automatic recognizer of three-dimensional objects appear to be greater than those of the fisheries laboratories, development should be undertaken to meet the needs of both.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

SECTION 1

INTRODUCTION

1.1 Background

The work reported here began after completion of operating models of the three-dimensional modelling process of binocular vision(1)* and of the core of the vertebrate central nervous system(2,3). These efforts, much larger than the one reported here, took place 3 to 6 years earlier. A subsequent study of means of automatically determining the age of fish(15) introduced the staff of the Draper Lab to problems of the fisheries. In the years since we made the models described in Refs. 1 and 3, understanding of human vision has advanced and components have so improved that we now dare propose a model of human vision as an instrument for use in a fisheries laboratory.

The remainder of this section describes the problem to which this study was addressed. Section 2 begins by considering methods of solving this problem other than by human vision. It then describes how human operators perform. Section 3 considers how a model of human vision may be devised. Section 4 first treats a detail of that model, namely, a binocular TV camera assembly, then concludes by describing how this assembly may be employed by a computer to classify objects such as plankton.

1.2 Need for Automatic Classification of Plankton

A problem faced by the National Maine Fisheries Service and research groups in fisheries is how to meet the needs of the fisheries industry for the classification of plankton. While the term plankton includes all living matter that drift with the sea, the fisheries industry is concerned mainly with the animal life, particularly the plankton that will mature into marketable fish. The study reported here was directed toward the automatic classification of the small animals called zooplankton.

An example of the commercial importance of classifying plankton is the following. A sample of sea water is collected at regular intervals from each 800 square mile area of the California current where anchovy have been caught. This sample is examined in the manner described below to determine, among other information, the quantity of anchovy larvae. From this quantity, an estimate is made of the anchovy eggs that have been laid and, from this the number of adult

---

*Superscript numerals refer to similarly numbered references in the List of References.

anchovy in the area sampled. This is called "backcasting." From this estimate, Dr. Paul Smith told us, the chicken food industry decides whether to supply its customers by anchovy ground into meal or by grain enriched with amino acids.(4,5)

Plankton form the bottom of the chain of sea life. Monitoring of this key to sea life is needed all over the world, but today it is done only in laboratories where skilled people have been especially trained for the task. Automated analysis equipment would be invaluable for providing rapid analysis of larger samples, establishing a broader, more reliable data base for evaluating ecological trends and providing more extensive, rapid monitoring of the vital life chain of the oceans.

1.3 What is Zooplankton?

Plankton is all the living matter that drifts with the sea. Since others are looking into the classification of vegetable plankton (phytoplankton)(6) the study reported here concerned itself only with the classification of animal plankton or "zooplankton".

Figure 1 shows how zooplankton interact as they age.(7) The herring shown getting larger across the top of the figure are plankton until they are strong enough to swim against the current. Note that the smallest herring are eaten by other plankton, as indicated by the dashed lines. An adult herring, in turn, eats a form of plankton (sagitta) that might have eaten the herring when the herring was younger.

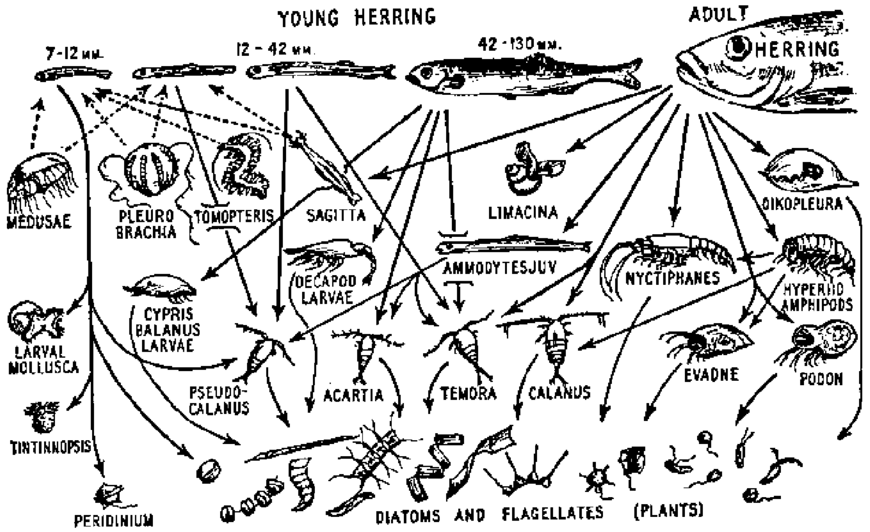


Fig. 1. A sketch showing the feeding relationships between the herring of different ages and the members of the plankton.

SECTION 2

STUDY OF MEANS OF CLASSIFYING ZOOPLANKTON

2.1 Possible Methods of Classification

A survey of opinion at the start of this effort indicated two principal possible approaches to automatic classification of zooplankton. One was chemical. The other, in imitation of the method employed by human classifiers, would use a TV camera, computer and manipulator. The chemical approach breaks down into three possibilities. One is to examine the DNA in each animal. A second is to grind up the animal, pyrolize it, form a gas chromatograph (GC) and examine the GC in a spectrometer. A third chemical possibility is to dissect the animal, analyze the contents of its stomach, gut and liver to determine what the animal ate and finally correlate this with the known eating habits of different species.

The following subsections treat each of these possible chemical methods of classification. Then we turn our attention to how human classifiers perform their task and how an automatic system can be devised to imitate their performance.

2.2 Classification by DNA

". . . the most fundamental feature of all living organisms is the capacity to produce faithful replicas of their specific proteins; only in this way can each organism perpetuate its identity. Control of protein replication in modern living cells is achieved through DNA, which contains the information for the amino acid sequences in the proteins of the organism and also for regulating their abundance. Consequently, the appearance of life on Earth represents the emergence of processes for copying polymers (proteins, nucleic acids). In addition, the copying mechanism also allows for an evolutionary process through which successive copies of polymers can be modified to suit the environment. Hart (1967) has recently proposed a detailed scheme for the emergence of a self-replicating system under primitive Earth conditions which also accounts for the appearance of nucleic acid control of protein synthesis subsequent to the synthesis of protein molecules themselves. Figure 2 includes a simplified version of Hart's hypothesis as part of a comprehensive summary of stages between the formation of the primitive atmosphere of Earth and the emergence of a stable self-replicating system. Hart recognizes six stages in the evolution of the replication process. . ." In the third

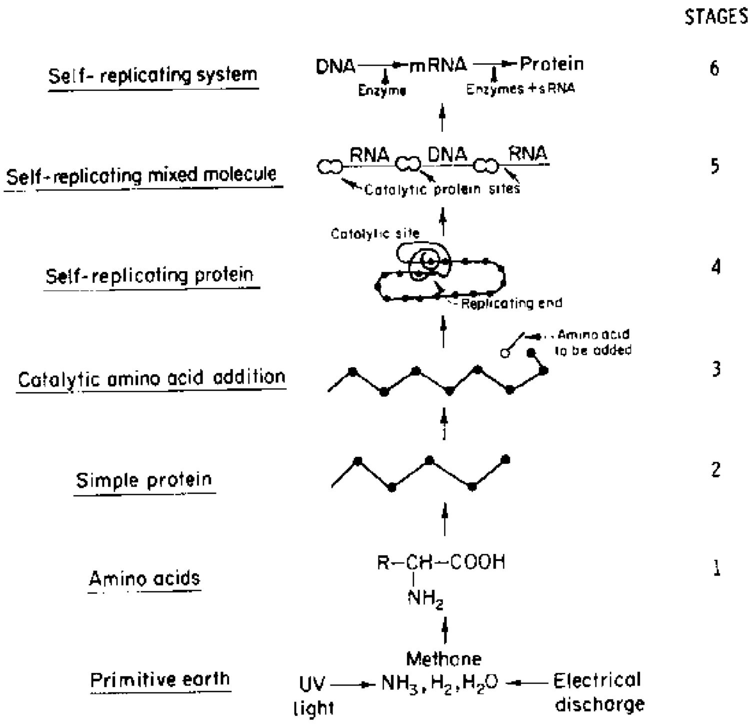


Figure 2. Hypothetical scheme for the emergence of a self-replicating system under primitive Earth conditions.

stage "the replicate. . . developed its own catalytic site, split off from the original chain, and began to transcribe itself. At this point, we have a system for self-replication of proteins on protein templates. The fourth stage of development was reached when the self-replicating molecule consisted not only of amino acids but also included components of nucleic acids which were transcribed by the catalytic site in the molecule. Here, the relationship of nucleic acid to protein synthesis was introduced into prebiological evolution. . . By now (step five) we have reached a recognizable precursor of the sequence DNA RNA protein seen in living organisms. Subsequent stages in Hart's scheme involve localization of the self-replicating units in a gel, in order to conserve the precursors and products of reaction and also to give the system the fundamental properties of life, namely to replicate, to have metabolism independent of the environment, and to influence the environment to the advantage of the organism."(8)

For each species of animal there is a unique DNA molecule which determines the form of all of the protein molecules in the animal. Identification of the DNA molecule can be performed by x-ray diffraction, but this process costs thousands of dollars and takes days to perform.(9) Each protein molecule is unique both to the species of animal and to the function it performs within the animal.

The following is an estimate of the proportion by weight of the principal types of molecules in one group of vertebrates, namely, the mammal.(10)

water	70%
protein	16
bone and cartilege	9
fat	5
other	2

Zooplankton have a larger proportion of bone and cartilege. This proportion varies widely with the species of zooplankton.

2.3 Classification by Grinding, Pyrolysis, GC and MS

The next question we asked was: Can the chemical components be identified with enough specificity to identify the zooplankton.

Prof. Ronald Hites (MIT) told us of a paper by P.G. Simmonds of Jet Propulsion Laboratory reporting the identification of two microorganisms, one of them a common soil bacterium, in the following steps(11):

1) 2 mg of dry whole organisms were placed in a stainless tube (inner diameter 0.127 cm; length 3 cm)
2) the sample was then pyrolyzed by heating to 500 C in 15 seconds in a helium atmosphere
3) the pyrolysate was separated on a capillary gas chromatographic (GC) column. The column was temperature programmed from 50 to 200 C at a heating rate of 4 deg/min.
4) mass spectra of the eluted components were recorded with an Electronic Associates Inc. quadrupole 300 mass spectrometer (MS). Compounds were identified by comparison of their mass spectral fragmentation patterns with library reference spectra.

Prof. Hites commented that variation in the chromatographic patterns between species is small, while variation between results obtained by different laboratories is as large as this variation between species; and the variation between runs in the same laboratory can be this large also. Furthermore, he added, to prepare for the above test a culture has to be grown. He thought the approach we describe in Section 3 more promising.(12)

When asked whether a larger animal, for example a plankton, might be identified automatically by grinding it up, pyrolyzing a sample, taking a GC of the sample then an MS, Dr. Max Blumer of the Woods Hole Oceanographic Institute replied that it is a "remote scientific possibility".

Dr. Blumer has analyzed the hydrocarbons of zooplankton in the stomach, gut and liver of a shark to learn from these analyses what the animal ate both recently and in the past. From these analyses he estimated the path of migration of the shark.(13)

Both the method of chemical classification described at the beginning of this subsection and the early development of a method of simulating visual recognition described in Sections 3 and 4 were developed under NASA support for the exploration of Mars.

2.4 Other Possible Methods of Automatic Classification

Figure 3 shows along its top line the existing technologies and techniques investigated in this study. At the lower right is a box labelled "computer engineering" into which lines from each of these technologies or techniques feed. It was recognized from the start that whatever data was collected by the methods in the top row would probably need to be processed by a computer. Moreover, the computer might be called upon to direct the test.

The possible chemical methods have been described above.

Ultraviolet radiation is employed in a device called a Fluorometer which causes chlorophyl to fluoresce and thus indicate the percent of phytoplankton in a sample.(14) This is a gross indication with no differentiation among species. To

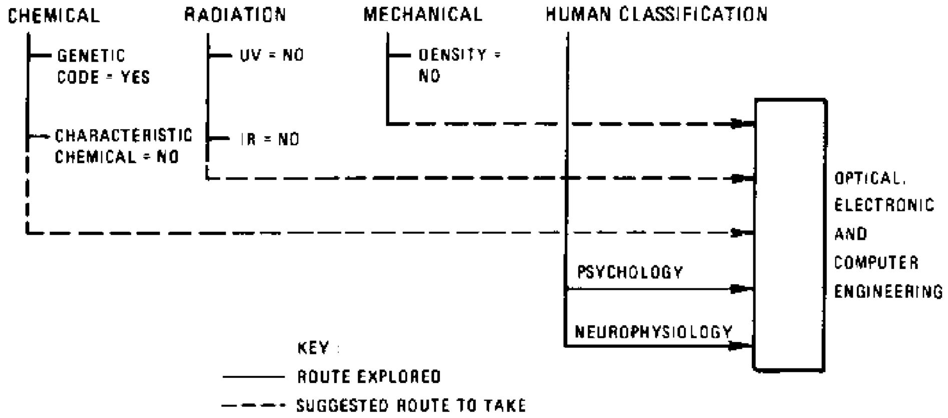


Fig. 3. Routes explored in this study.

perform the differentiation, a unique chemical has to be discovered for each species of phytoplankton as well as of zooplankton.

Infrared radiation (IR) is another form of the heat whose application we examined in subsection 2.3.

A mechanical means has been developed of separating a sample of plankton into "fish eggs, fish larvae and invertebrates with only modest overlap among these three major types."(6) The method was demonstrated first in a centrifuge, then was tried in a tall glass bottle. The latter works very well. Not only do the three major types settle into different height zones in the bottle, but different species of phytoplankton settle at different heights in the bottle. Unfortunately, species of fish larvae do not sort themselves out this way.

2.5 Human Vision

The only technique employed today to classify plankton is human vision, which scientists describe in different ways. The late Dr. Warren McCulloch described vision in terms of relations. For example, what appear to be rings around the the bodies of the anchovy* and sardine** in Fig. 4 are the boundaries between muscle segments or metameres. (A metamere is used as a unit of relative distance in visual classification of zooplankton as explained below.) Each ring appears to encircle the plankton and the rings are approximately parallel. Both of these last two statements are relations.

A fin is defined by more relations: It is very thin compared to the plankton's body. It protrudes from the smooth contour of that body.

Differences in relations can be used to distinguish one species of plankton from another. In Fig. 4 the two right-hand circles on the anchovy show that in that plankton the beginning of the anal fin is directly beneath the end of the dorsal fin. The corresponding circles on the sardine show that the anal fin begins 5 to 8 metameres beyond the end of the dorsal fin.(16)

The forward circles in Fig. 4 enclose dashes of pigment which extend along the gut tube to the anus in both species but are longer in anchovy than in sardine. "The gut or digestive tract is almost a tube in anchovy and sardines, becoming wider at its posterior end."(16) In larvae smaller than those shown in Fig. 4, the gut alone can be used as a measure. "In anchovy the gut is 70% of the full length; in sardines 80 - 85%".(16) A binocular TV-camera computer should be able to locate this tube and the anus at its end, and measure its length.

Relations can be composed of simpler relations. To quote Dr. Irvin Rock "A prevailing view among psychologists and sensory physiologists is that form

---

*Sardinops caerulae

**Engraulis mordax

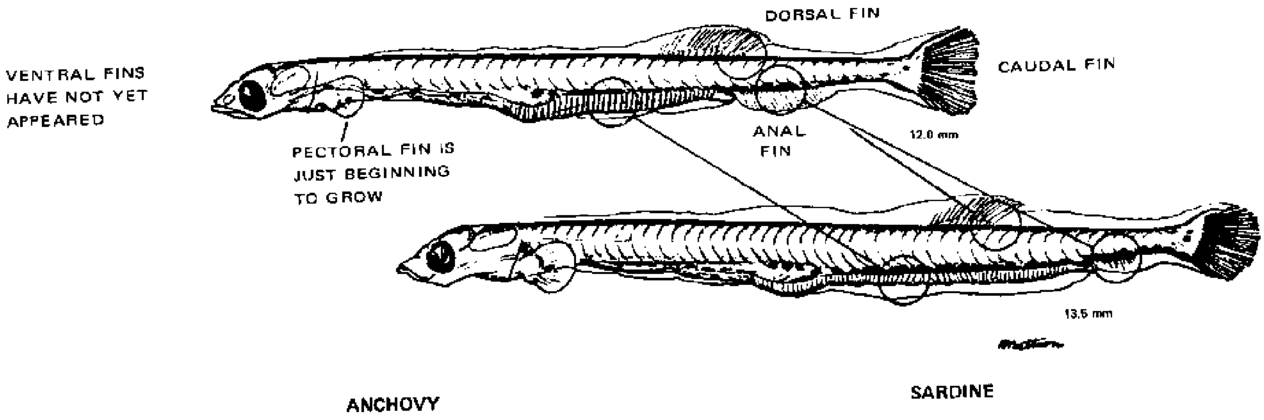


Fig. 4. Features (circled) by which an anchovy plankton can be distinguished from a sardine plankton (from instruction sheet used at Southwest Fisheries Center, National Marine Fisheries Service, LaJolla, Cal.).

perception can be reduced to the perception of contours and that contour perception in turn can be reduced to abrupt differences in light intensity that cause certain neural units in the retina and brain to fire. . . Such an explanation is far from sufficient. . . The perception of form depends on certain mental processes such as description and correction".(17) Section 3 presents a method of description employed in human vision namely, scan paths of successive fixations. What Dr. Rock means by "correction" is rotation, which can be avoided if the specimen is placed in a standard position, such as that of Fig. 4, before recognition is attempted.

The processing which causes "certain neural units in the brain to fire" and which Dr. Rock calls "not sufficient" is nevertheless necessary to vision. The cells of which we speak have been investigated by Hubel and Wiesel(33, 34, 35). We plan to combine the results of this work with the work of Noton and Stark which we describe in Section 3.

2.6 Human Classification of Zooplankton

The author spent a day watching an expert identifier sorting plankton at the National Marine Fisheries Service, Southwest Fisheries Center in La Jolla, Cal. Although not a college graduate she had studied the zoology of the plankton and was completing her seventh year of sorting. Few people, we understand, are willing to study this subject as she has or stay with this kind of work as long. Now she has left that laboratory.

The features she and others looked for in distinguishing a sardine plankton from an anchovy plankton, both of approximately 1/2 in. length, are shown circled in Fig. 4. That illustration is one of many provided by that laboratory to its identifiers and sorters. The author also watched sorting at the Narragansett Fisheries Laboratory and at the Woods Hole Oceanographic Institute. On the basis of these observations, and of references given below, he concludes:

(1) Manipulation of a plankton with forceps is necessary to lay it out so that it can be viewed as in Fig. 4.

(2) Only the features circled in Fig. 4 and others needed for orienting the specimen need to be examined in detail.

(3) This examination of detail is called by the psychologist Trevarthen, "focal vision".(18) Low-resolution vision to view the outline of the specimen and its orientation he calls "orienting vision."

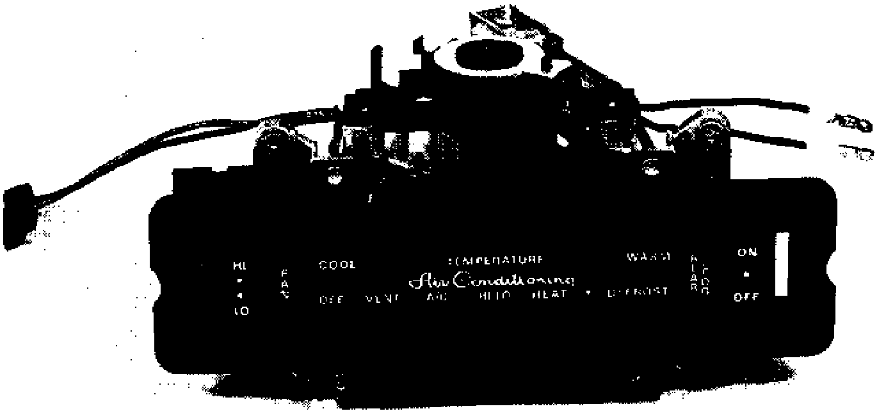
(4) Vision needs to be binocular so that a feature will appear the same shape from whatever direction it is viewed.

(5) Recognition consists of seeking a previously stored sequence of features.(19,20)

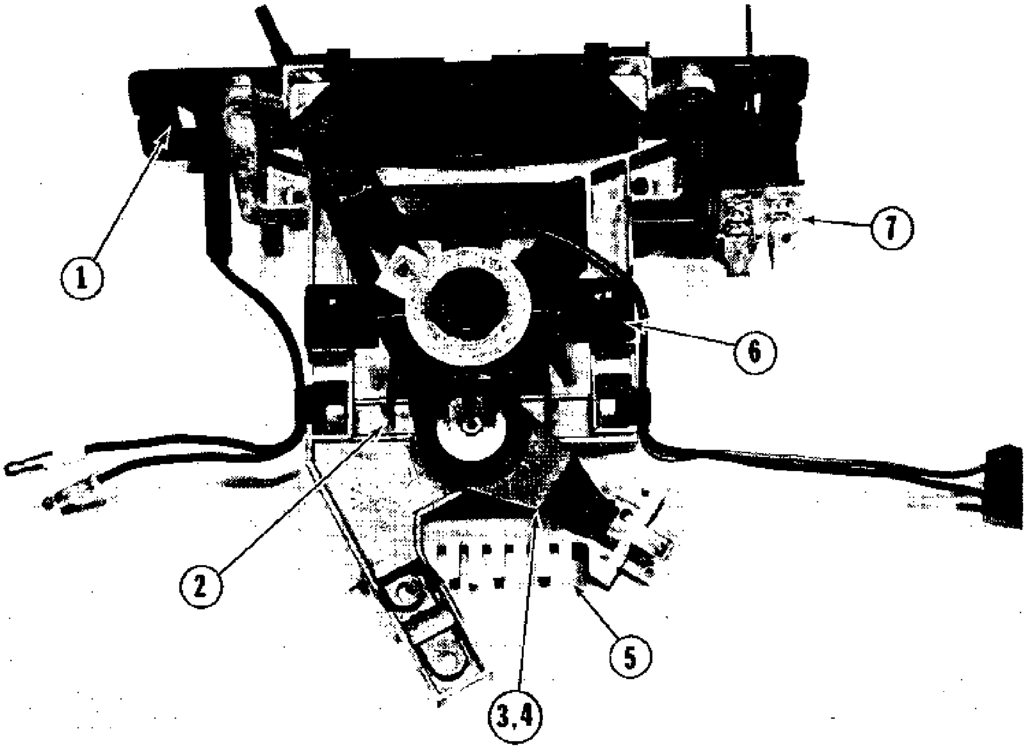
2.7 Other Examples of Human Classification

An area where human classification is used on a large scale and where there is the greatest pressure to replace the human operator by a automatic device is in industry. There the problem is called inspection and consists of classifying manufactured parts and assemblies by their defects.

An example that we have studied for the Ford Motor Co. is the heater and air conditioning control assembly pictured in Fig. 5. This assembly was used in every 1973 Ford automobile. Defects that appeared sufficiently frequently to make inspection desirable are the seven features numbered in circles. These are features on which a trained human inspector fixates successively, performing less intensive inspection between these fixations.



Front view



View from above rear showing the following possible defects:

1. Wrong face plate can be installed as indicated by the lack of this aperture.
2. Pointed tip can be missing, indicating wrong part.
3. Pin can be missing that links arm labelled "temperature" on front panel to this arm so that this arm does not follow "temperature" arm.
4. This arm can be missing.
5. Wrong 8-port (of which 7 ports are shown) can be installed.
6. Screw can be missing.
7. There can be a wrong blower switch, as indicated by the wrong number or wrong orientation of terminals.

Fig. 5. Ford heater and air conditioning control assembly.

SECTION 3

THE CYBERNETIC APPROACH; POSSIBILITY OF BUILDING A PRACTICAL SYSTEM

3.1 The Cybernetic Approach

In 1948, Norbert Wiener startled the scientific world with a book that used the same logic and mathematics to describe both animals and machines. The book, "Cybernetics", showed that the control system known as a servomechanism is necessary to the operation of both animals and automatic machinery.(21) The book also described the contributors to this newly labelled field. Prominent among them was a neurophysiologist who thought like an engineer, Dr. Warren McCulloch. We say "thought like an engineer" because his diagrams and descriptions of the nervous system often look more as if they had been designed than as if they had been discovered.

McCulloch enlarged the realm of commonality among animals and machines to include what was then called "computation" but is more often today called "information processing". McCulloch's point was that there are realms of the vertebrate central nervous system which could be called "special" computers. He used the word "special" to mean "intended for a special purpose" as contrasted to the general-purpose computer commonly used today.

Cybernetics works two ways. Not only does it describe the vertebrate central nervous system in engineering terms but it also suggests that the equivalent of that system can be built. For Ref. 2, McCullouch drew a block diagram of the vertebrate central nervous system containing five special computers. Reference 3 describes an operating model we have devised of one of these, namely, the core of the central nervous system in the reticular formation. Reference 15 describes a computer to first detect edges in the manner of the visual cortex of monkeys and man, then locate these edges in three-dimensional space.

3.2 The Recognition of Objects

The scheme by which a human being recognizes (or classifies) an object has been described by Noton and Stark as follows:

1. "The internal representation or memory of an object is a piecemeal affair: an assemblage of features or, more strictly of

memory traces of features; during recognition the internal representation is matched serially with the object, feature by feature.

2. The features of an object are the parts of it (such as the angles and sharp curves of line drawings) that yield the most information.
3. "The memory traces recording the features are assembled into the complete internal representation by being connected by other memory traces that record the shifts of attention required to pass from feature to feature, either with eye movements or with internal shifts of attention; the attention shifts connect the features in a preferred order, forming a feature ring and resulting in a scan path, which is usually followed when verifying the features during recognition."

Subjects were asked to look at Fig. 6a which was made sufficiently large and placed sufficiently close to the eyes that it could not be registered on the fovea in one fixation. The other parts of Fig. 6 shows " [the] recurrence of [the] scan path during recognition of an object. . . [as] predicted by the feature-ring hypothesis. A subject viewed at (a) an adaptation of a drawing by Paul Klee. A scan path appeared while he was familiarizing himself with the picture (b, c). It also appeared (d, e) during the recognition phase each time he identified the picture as he viewed a sequence of familiar and unfamiliar scenes depicted in similar drawings. This experimental subject's scan path for this particular picture is presented in ideallized form at f."(22) Each number in that figure marks a fixation of the eyes, each line between numbers a rapid flick of the eyeballs called a "saccade".

Features were examined in the following sequence:

A. Base of nose
B. Forefinger of right hand
C. Little finger of left hand
D. Left eye
E. Back of head

Examples are given in Ref. 22 of drawings of an automobile and a tree and an abstract drawing, all recognized by similar scanpaths of successive fixations. Reference 16 reports that "the appropriate scan path appeared in about 65 percent of the recognition-phase viewings". The authors consider this "a rather strong result in view of the many possible scan paths around each picture". They suggest that the "feature rings" consisting of fixations and the eye motions between fixations, which they hypothesized to be the memory trace, could have within it other

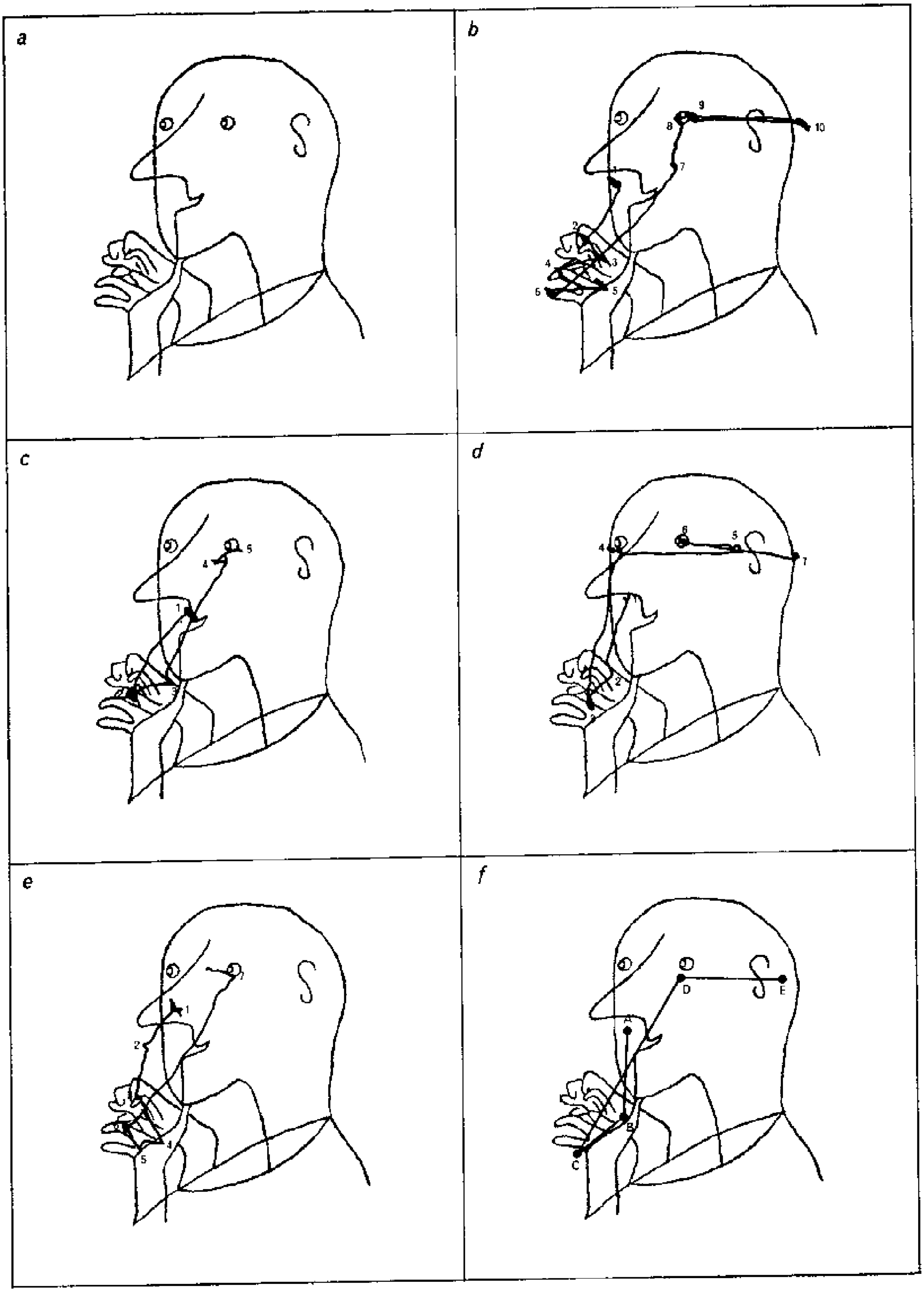


Fig. 6 Recurrence of scan path during recognition of object. (From "Eye Movements and Visual Perception" by David Noton and Lawrence Stark. Copyright © by Scientific American, Inc. All rights reserved.)

eye movements between features not adjacent in the ring. The original feature ring would then represent the preferred and habitual order of processing rather than the inevitable order.

We propose to employ a scan path such as that in (f) for the automatic recognition of plankton. That is, the two optical axes of the camera assembly would flick from feature to feature.

3.3 Need for Both Central and Peripheral Views

In both the study made for NASA and the present study two lines of thinking led to a binocular TV camera with central and peripheral views. One path is the necessity of providing a means, within an automatic system, of first locating each feature that needs to be identified in the classification of objects, and second of identifying that feature regardless of its three-dimensional orientation. The second line of thinking is the evidence that monkeys and men have both the locating and identifying vision we need in a machine.(23) In these primates, the input to locating vision is the periphery of each retina, the input to locating vision in the fovea. The angle subtended by the fovea is 2° horizontally and 1° vertically.(13) The angle subtended by the periphery is approximately 180° in primates.

3.4 Binocular Vision

To employ a scanpath of the kind described in 3.2, each feature will be detected in the peripheral fields of a binocular TV camera, then the central fields will be aimed directly at that feature. A binocular TV camera and computer, of the kind we are presently operating, forms a three dimensional model of the feature in its memory so that the pattern of the feature can be recognized no matter how the feature is oriented.

3.5 A Self-Command Computer

Vision is purposeful. It assists an animal to explore, hunt, escape or pursue a dozen or more other purposes. To determine this purpose from moment to moment we have devised a self-command computer.(3) To explore a plankton, the system we consider here will position it. To hunt for its features, its optical axes will follow a scanpath suggested by the first feature identified. After classifying a plankton, the system will escape from this task by discarding the plankton. While a computer as intricate as that of Ref. 3 may not be needed, it helps our planning to know that the design of such a computer is available.

3.6 Possibility of Building a Practical System

Section 4 presents a preliminary design of a binocular TV camera, with central and peripheral views to serve as the input to an automatic recognizer or classifier. However, the largest part of this automatic recognizer will be an assembly of the kinds of computers McCulloch observed in the vertebrate central nervous system.

Can computers be built that can perform, with the camera assembly described in Section 4, recognition equivalent to or better than that performed by human beings? The approach we have taken to answering this question is to imitate the spirit if not the fine details of each equivalent human computer.

Can the desired computers be manufactured for a price that the Fisheries Service could afford? We can answer this question only in terms of trends in the cost of computers and in the demand for automatic equipment.

The first trend is toward smaller and smaller electronic components at less and less cost. "Cost of a digital gate* — the building block of the digital computer — constitutes the most significant factor driving the advances since 1950. In 1950 a gate in vacuum-tube technology cost $20 - $40; today a gate in integrated-circuit technology costs an order of magnitude less; and by 1980 a typical gate should cost 2 - 5 cents: A factor of 1000 cost decrease in 30 years — at the same time gate reliability has increased roughly three orders of magnitude and physical size has decreased the same!"(24) Components of TV cameras are also getting smaller and more reliable as we explain in Section 4.

3.7 Cybernetic Design

Electronic components are approaching the size of neurons (see Subsection 4.2) but the number of connections to each component is only a few, while the number of connections to a neuron may be in the thousands. An electronic component, on the other hand, operates in the order of a million times faster than a neuron. By time sharing electronic components it is becoming possible to achieve the connectivity of the nervous system. Digital electronic components include means of storing a number anywhere in the electronic system. Thus intermediate sums, products, etc., can be formed and operated on at irregular intervals.

3.8 Direction of the Proposal

A proposal to continue the work reported here is being prepared for submission to the federal government, foundations and industry. The proposal aims to meet first the needs of industry for automatic inspection of three-dimensional objects The needs of industry are great enough, we believe, to invite the investment of capital and lead to the manufacture of the kind of equipment that can then be used by the National Marine Fisheries Service.

---

*Two gates can be connected to form a flip flop, a basic memory element of a computer.

SECTION 4

BINOCULAR TV CAMERAS WITH BOTH CENTRAL AND PERIPHERAL VIEWS, OPERATION UNDER COMPUTER CONTROL

4.1 Introduction

The work reported in this section is a continuation of a study, begun for the National Aeronautics and Space Administration (NASA), of a TV camera assembly to serve as the front end of a system to automatically recognize objects. The system studied for NASA was intended to explore the surface of Mars and to report back what it saw. The problem reported on here differs not only in scale, but also in that classification is among a limited number of species and the system is stationary.

4.2 Sensor Arrays

Reference 22 describes a succession of binocular TV camera assemblies that we have designed and some of which we have built. One design in that series provides both central and peripheral views on each binocular axis. Because those camera assemblies were designed three to six years ago a vidicon camera tube was employed to transduce each image to an electric signal.

Now there are four possible transducers besides camera tubes. One is the new 100 x 100 = 10,000 element solid-state array made by Fairchild Camera Instrument Co. A second possiblity is an array of 80,000 sensing elements which would come "pretty close to duplicating a standard. . . vidicon". RCA, whose market planner Minet made this statement, has built and operated a 60,000 element array. "RCA's goal is to produce devices compatible with 525-line TV receivers, points out Karl H. Zaininger, head of solid-state-device technology at RCA's Sarnoff Research Center, Princeton, N.J." Such an array of approximately 500 x 500 elements -- our third possibility -- will permit low-light level applications* such as we need for the beam splitting we describe below. "Two to three years from now" is Minet's estimate for this item.(25)

A fourth possibility is the development of a solid-state array whose sensing elements are as closely spaced as the cells in the fovea of the human eye. We propose to proceed in a way that will exploit each of these arrays as it becomes available.

---

*because resolution should not decline with illumination as it does in a vidicon

A binocular TV camera assembly, designed for NASA, is shown in Fig. 7a. Like each human eye, each camera yaws. Unlike human eyes the two cameras pitch together and they do not roll. However, all of the motions of the head are there: pitch, roll and yaw (azimuth). Figure 7b is a vertical section through the camera showing how approximately 1% of the light entering at the left is reflected by a beam splitter upward onto the face of the vidicon, forming there the low-resolution 20° field shown in Fig. 7c. Approximately 99% of the light in the central 2° of the field of view is magnified 10X by a microscope objective (Fig. 7d) which forms the high resolution central field shown in Fig. 7c. The light is divided in a ratio of approximately 100 to 1 because the area magnification of the narrow-angle view with respect to the wide-angle view is 100 to 1. Dividing the light this way assures approximately even illumination of the vidicon.

4.3 Gimbals and Vidicons

While we have not built the assembly of Fig. 7 we have built a simpler binocular assembly$^{(24)}$. From experiments with the latter we conclude that there is a trade off between accuracy of positioning the binocular axes and computation. That is, the positioning can be cruder than is shown in Fig. 7 and compensated for by corrections made in the interpretation of the images. Development of such information processing will be proposed.

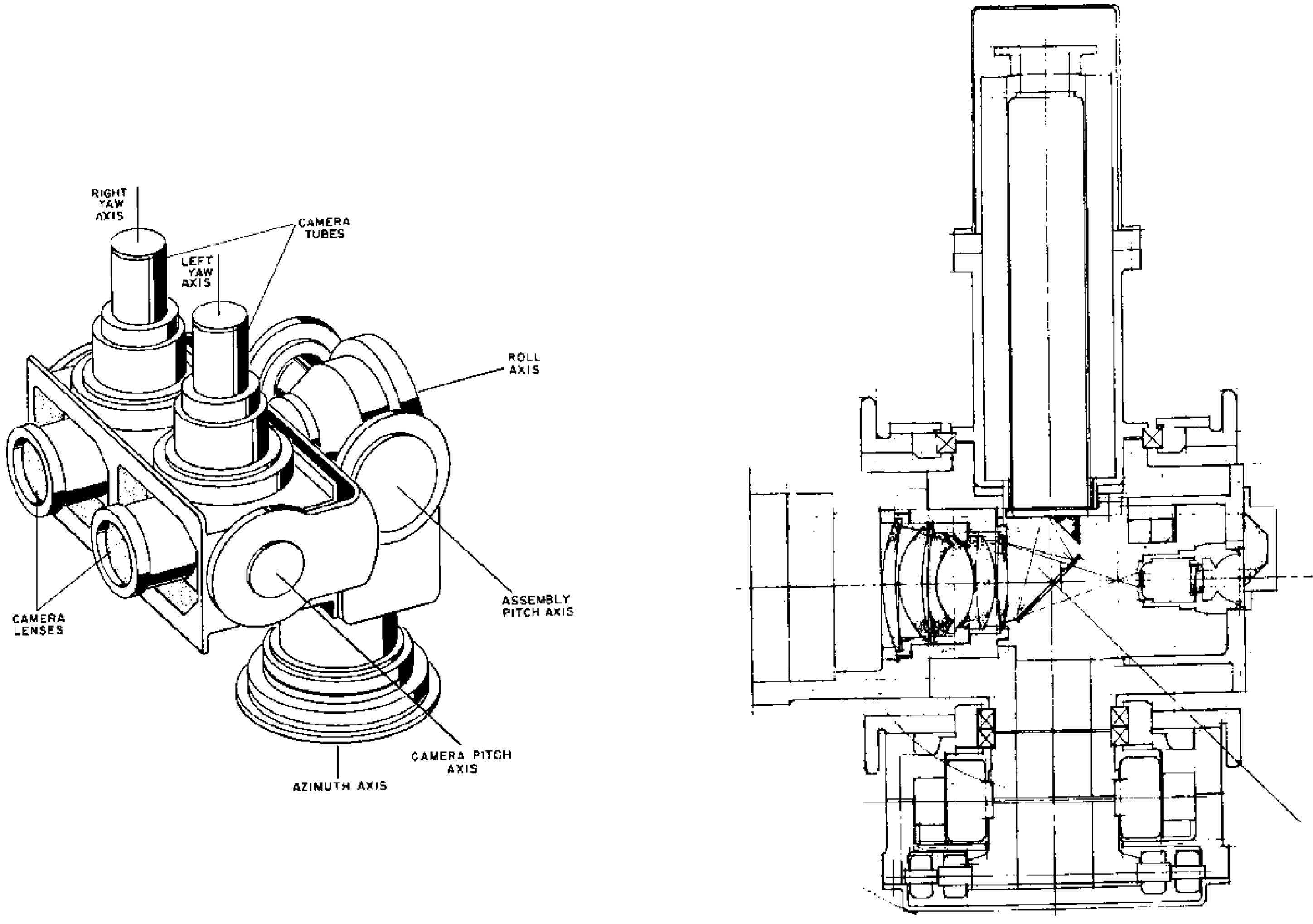
For automatic positioning of a plankton under a microscope and classification of that plankton by viewing through the microscope all of the degrees of freedom shown in Fig. 7 except the rear pitch gimbal are needed. The degrees of freedom of each camera might be provided by making each camera a sphere and turning it in a socket as the human eye is turned. The electron-fabricated sensor array described in 4.5 could make this possible.

4.4 Solid State Sensors and Liquid Prisms

Between the design of Fig. 7 and an eyeball-like camera are many useful steps. Four steps can be envisioned, each employing one of the four possible forms of sensor arrays described above.

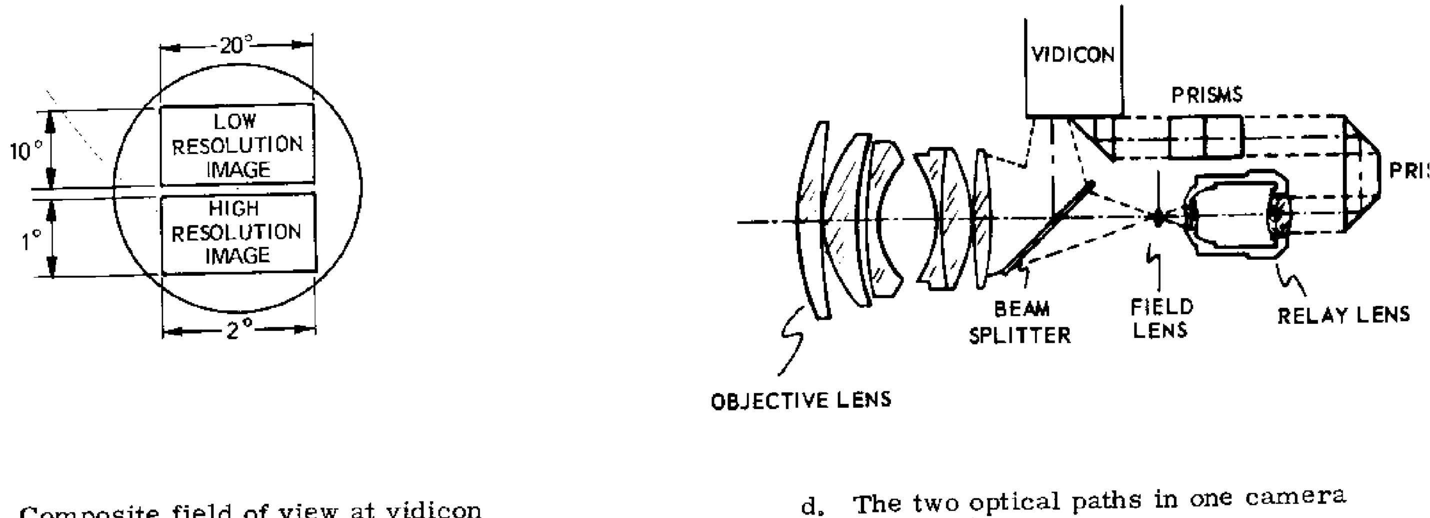
The configuration of Fig. 8 is intended to provide, with the third type of solid-state array, one-half the resolution of the human eye (see 4.5). While we are awaiting the third type of array we can experiment with the first and second. The lower part of Fig. 8 diagrams a conventional binocular microscope, but for two modifications. The objective above the plankton is one of the usual set that provides, together with the eyepiece, a range of magnifications from 10X to 60X.

The first modification is the replacement of standard eyepieces by lenses that form images.



a. Outside view

b. Section through one camera



c. Composite field of view at vidicon

d. The two optical paths in one camera (see text)

Fig. 7. Binocular TV camera assembly using vidicons as sensors and gimbals for camera motion.

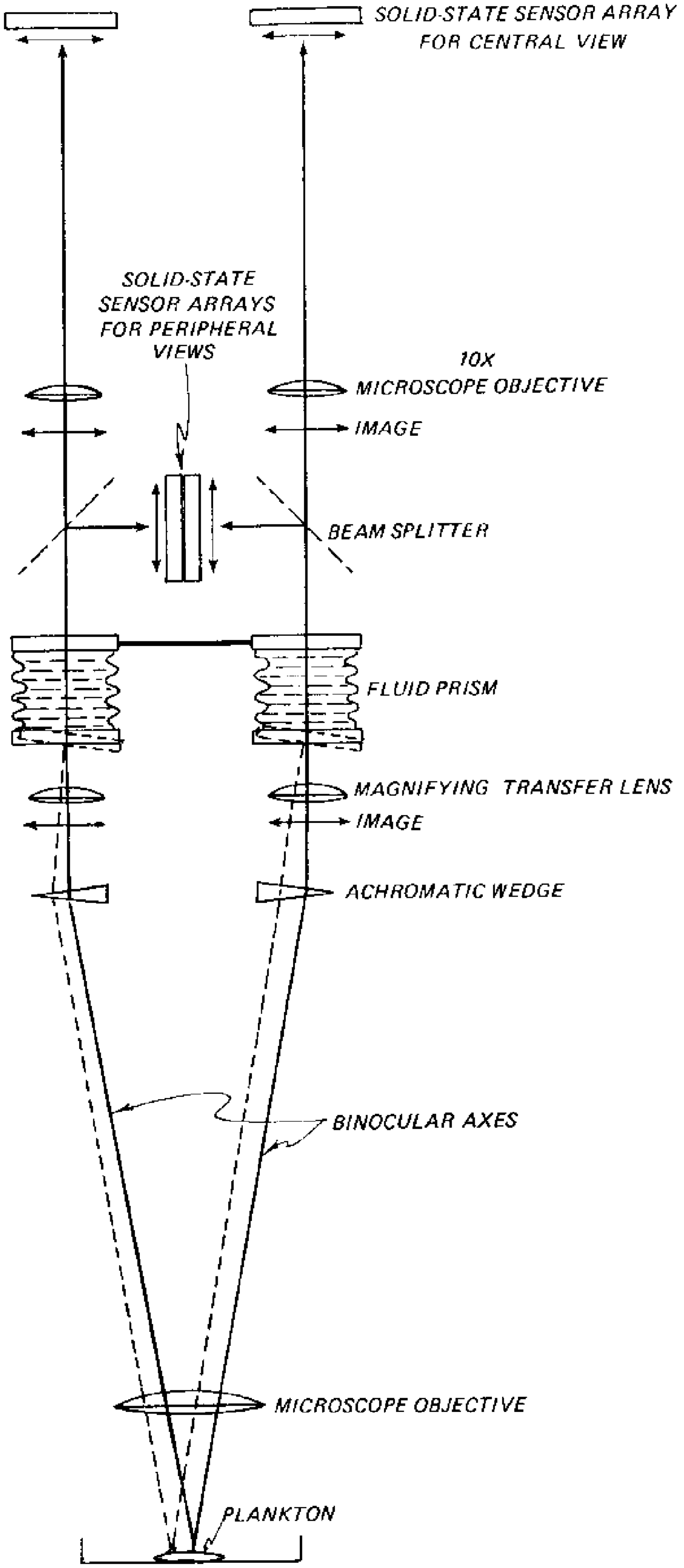


Fig. 8. Combined binocular TV camera and microscope.

The remainder of the assembly of Fig. 8 models human eyeballs. Liquid prisms manufactured by Dynasiences Corp., Blue Bell, Pa., provide two degrees of freedom. One degree of freedom is that of the two axes in their common plane. The dashed lines show the effect of turning the front plates of the liquid prisms to shift the optical axes to a different part of the plankton. The second degree of freedom is provided by tilting the rear plates, shown here ganged together. A peripheral view of plankton features is obtained by the lower pair of sensor arrays. A central view, of one feature at a time, is obtained by the upper pair of sensor arrays. The central view is rigidly fixed within the peripheral view.

4.5 The Possibility of Matching the Resolution of the Human Eye

The separation of cones at the center of the fovea of the human eye is 2 to 3 microns or 10^{-4} in.(27) Using electron fabrication "an experimental transistor structure" has been made "with aluminum lines 0.5 micron wide separated by 0.5 micron gap."(28)

The pattern of nearly every integrated circuit today is impressed on a photoresist by shining ultraviolet light on it through a mask outlining the desired pattern. An electron beam if used for this purpose can achieve the ten times greater resolution represented by the dimension of the transistor given above. At a three-day conference on "Electron, Ion and Laser Beam Technology" held at MIT last spring, one of the participants told us that his company is experimenting with arrays of sensors of those dimensions. We have been reliably informed that several companies are preparing to manufacture large-scale integrated circuit by electron fabrication.

Thus if the assembly of Fig. 8 is built it can be expected to be tested with arrays that can achieve the resolution of the human eye — or better — without the 10X magnifier shown in Fig. 8. Not only can the lens be removed but the beam splitter also. A single array behind each liquid prism will suffice. With the removal of the beam splitter the need for sensitivity will be reduced by a factor of 10.

The resolution of the fovea accounts only for what Julesz calls "ordinary visual acuity",(27) which is well correlated with the anatomical spacing of retinal cells given above. In addition there is a superresolution called "vernier acuity" which is ten times the anatomical resolution. In Ref. 30, Platt reasons that this increase is due to both the random pattern of cells in the retina and the motion of the eyeball across, say, a black line, on a luminous surround, whose gap is to be detected. Lack of stimulation from the two ends of the black line may be compared by the brain to stimulation from the gap between the ends as the eyeball sweeps across the line. In his theory, all that are needed to detect this gap are two cells not stimulated by the ends of the line and one stimulated by the luminous gap. A random pattern of cells makes it more likely that three cells will be in the right places to achieve maximum acuity.

Such random distribution of sensory elements could be achieved by electron fabrication.

4.6 Operating A Binocular TV Camera Assembly with A Computer

To test the usefulness of a binocular TV camera assembly, it needs to be operated as part of a model of vision. The model C3 assembly has been operated to acquire left and right images of a scene which the processing described in Ref. 1 can fuse into a three-dimensional model of that scene. The next step is for the computer to direct this assembly to search a scene for desired features and, when one is found, attempt to identify it as part of the scan-path of an object.

A graduate student in electrical engineering at MIT and a mechanical engineer, volunteering his time, are completing the D1 assembly so that its binocular axes can be directed by a computer to scan a scene in search of each feature. That graduate student then plans to program the computer to locate three-dimensional features and identify them. This system, employing only one lens on each optical axis, anticipates the system of Fig. 8.

SECTION 5

CONCLUSION

There appears to be only one promising route to follow in the development of automatic classification of zooplankton. This is to apply the principles of human vision to the design of an opto-electronic computing device. A manipulator or manipulators may be part of this device.

A popular book has been written arguing that such a device cannot be made to work.(32) Why do I think it can?

For three reasons. One is that the human central nervous system can be represented as multiple loops of information flow with processing in every loop. The second is that this processing can be performed by increasingly small, fast and inexpensive components. The third is that the performance of human eyes is being approached and may soon be matched by devices under development.

The largest part of the required development is the coding of the relations which can tie edges into lines, lines into features and features into objects. The scale of the development is such that it should either follow another development of a comparable need, e.g., industrial inspection, or accompany it.

The development of a computer-based system analogous to vision, for the classification of zooplankton, appears to be achievable in the 1980's. Although results are encouraging to date, a substantial effort, in the order of several million dollars, will be required to carry the development through to an operating system.

REFERENCES

1. Sutro, L. L. and J. B. Lerman, "Robot Vision," Report R-635, Charles Stark Draper Laboratory, 1973.

2. Sutro, L. L., "The McCulloch Approach to Both Modelling the Vertebrate Central Nervous System and Designing Robots," Report R-636, Charles Stark Draper Laboratory, in preparation.

3. Sutro, L. L. and W. L. Kilmer, "A Self-Command Computer, Modelled after the Core of the Reticular Formation, with an Ability to Learn, Modelled after Classical Conditioning," Report R-736, Charles Stark Draper Laboratory, in preparation.

4. Conversation with Dr. Paul E. Smith, Director, National Marine Fisheries Service, Southwest Fisheries Center, LaJolla, Cal., September, 1972.

5. Smith, P. E., "The Increase in Spawning Biomass of Noethern Engraulis Mordax," Fishery Bulletin, Vol. 70, No. 3, 1972, National Marine Fisheries Service, Southwest Fisheries Center, LaJolla, Cal., 92037.

6. Price, C. A., "Analysis of Plankton Populations by Sedimentation in Density Gradients," Interim Report for the period 1 June 1972 to 14 March 1973, available only from the author who is director of Particle Separation Facility, Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, N. J.

7. Hardy, A., The Open Sea: Its Natural History, Part II, Fish and Fisheries, Houghton Mifflin Co., Boston, 1965, p. 62.

8. Munro, H., "An Introduction to Protein Metabolism during the Evolution and Development of Mammals," Mammalian Protein Metabolism, Vol. 3, Academic Press, N.Y., 1969, pp. 8, 9.

9. Holmes, K.C. and D. Blow, The Study of Protein and Nucleic Acid Structure, Interscience Pub., New York, 1966.

10. Munro, H., personal communication to the author, 1973.

11. Simmonds, P. G., "Whole Microorganisms Studied by Pyrolysis-Gas Chromatography-Mass Spectrometry: Significance for Extraterrestrial Life Detection Experiments," Applied Microbiology, Vol. 20, No. 4 (Oct. 1970), p. 567-572.

12. Hites, R., Dept. of Chemical Engineering, M. I. T., personal communication to the author, 1973.

13. Blumer, M., "Hydrocarbons in Digestive Tract and Liver of a Basking Shark," Science, Vol. 156, No. 3773, pp. 390-391.

14. Fisheries Board of Canada, "Some Current Problems in Marine Phytoplankton Activity," Technical Report No. 370.

15. Sutro, L. L., "Study of Automatic Means of Determining the Age of Fish," Report MITSG-2, Marine Resources Information Center, Room 5-331, M. I. T., Cambridge, Mass., 02139, 1971.

16. Allstrom, E. H., personal communication to the author, 1973.

17. Rock, I., "The Perception of Disoreinted Figures," Scientific American, January 1974, p. 85.

18. Trevarthen, C. B., "Two Mechanisms of Vision in Primates," Psychologishe Forshung, Vol. 31, No. 4, pp. 299-337, 1968.

19. Noton, D. and L. Stark, "Scanpaths in Saccadic Eye Movements While Viewing and Recognizing Patterns," Vision Research, Vol. 11, pp. 929-942.

20. Noton, D. and L. Stark, "Eye Movements and Visual Perception," Scientific American, June 1971, pp. 34-43.

21. Wiener, N. R., "Cybernetics," John Wiley & Sons, New York, 1948.

22. Ref. 20, pp. 35, 40, 41 and 43.

23. Held, R., et al., "Locating and Identifying: Two Modes of Visual Processing; A Symposium," Psychologische Forschung 31, 42/43, 1967.

24. List, B., "DAIS: A Major Crossroad in the Development of Avionic Systems," Astronautics & Aeronautics, January, 1973, p. 61.

25. "CCD image array is biggest ever," Electronics, September 27, 1973, pp. 33, 34.

26. Ref. 1, pp. 51, 52.

27. The Eye, edited by Hugh Davson, Vol. 2, The Visual Process, pp. 16-18, Academic Press, New York, 1962.

28. Broers, A. N. and A. M. Hatzakis, "Microcircuits by Electron Beam," Scientific American, Vol. 227, No. 5, November 1972, pp. 34-55.

29. Julesz, B., Foundations of Cyclopean Perception, University of Chicago Press, 1971, p. 78.

30. Platt, J. R., "How We See Straight Lines," Scientific American, June 1960, pp. 121-129.

31. Sutro, L. L., Notes on conversations with W. S. McCulloch, unpublished.

32. Dreyfus, H. L., What Computers Can't Do, Harper & Row, New York, 1972.

33. Hubel, D.H., and Wiesel, T.N., "Perceptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex, Journal of Physiology, Vol. 160, pp. 106-154, (1962).

34. Hubel, D.H., and Wiesel, T.N., "Perceptive Fields and Functional Architecture in Two Non-Striate Visual Areas of the Cat", Journal of Neurophysiology, Vol. 28, pp. 229-289, (1965).

35. Hubel, D.H., and Wiesel, T.N., "Perceptive Fields and Functional Architecture of Monkey Striate Cortex, Journal of Physiology, Vol. 195, pp. 215-243, (1968).