

UNCLASSIFIED

AD NUMBER
ADB193721
NEW LIMITATION CHANGE
TO Approved for public release, distribution unlimited
FROM Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; JAN 1962. Other requests shall be referred to National Aeronautics and Space Administration, Washington, DC.
AUTHORITY
NASA TR Server Website

THIS PAGE IS UNCLASSIFIED

ALLIED RESEARCH ASSOCIATES, INC.



BOSTON, MASSACHUSETTS

0

108.395

AD-B193 721



January 1962

COPY 1

Document No. ARA-1025

Prepared for

National Aeronautics and Space Administration
Office of Life Sciences
Washington 25, D. C.

LIBRARY COPY

DTIC
ELECTE
AUG 22 1994

S

G

D



SUMMARY REPORT ON A REVIEW OF
BIOLOGICAL MECHANISMS FOR APPLICATION
TO INSTRUMENT DESIGN

N-108,395

N-108,395

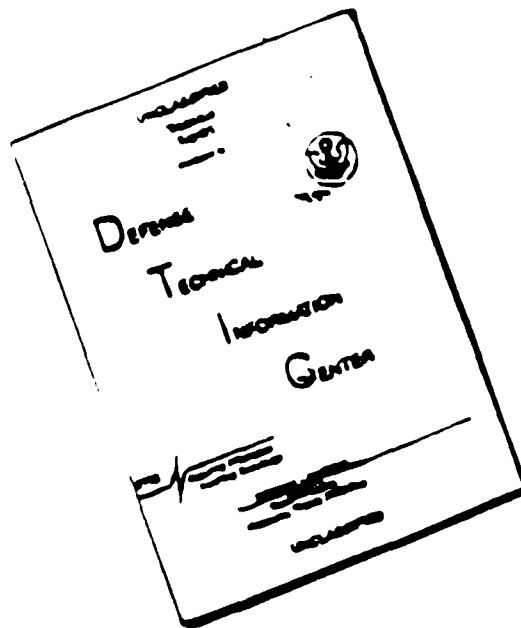
94-26486



20308

94 8 19 01 6

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST
QUALITY AVAILABLE. THE COPY
FURNISHED TO DTIC CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.



January 1962

SUMMARY REPORT ON A REVIEW OF
BIOLOGICAL MECHANISMS FOR APPLICATION
TO INSTRUMENT DESIGN

by

J. Healer

A. Arms

J. Merchant

E. Britt

Accession For	
NTIS CRA&I	☐
DTIC TAB	☒
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
12	

Prepared for

National Aeronautics and Space Administration
Office of Life Sciences
Washington 25, D.C.



FOREWORD

The study presented in this report was performed for the office of Life Science Programs, National Aeronautics and Space Administration under Contract No. Nasr - 16. The work was accomplished under the jurisdiction of Captain Frank E. Voris, MC USN, Assistant Director of Aerospace Medicine.

The program has been an interdisciplinary study of biological mechanisms, particularly biosensors, for application to instrumentation design and engineering.

TABLE OF CONTENTS

	<u>Page</u>
FOREWORD	iii
LIST OF FIGURES	vii
SECTION I INTRODUCTION	1-1
1.1 General Background	1-2
1.2 Scope of the Study	1-9
1.3 Exploitation of Biological Mechanisms	1-10
1.3.1 Direct Imitation of Biological Mechanisms	1-10
1.3.2 Utilization of Living Biosensors	1-11
1.3.3 Utilization of Basic Operational Principles of Biomechanisms	1-13
1.3.4 The Development of a New Technology	1-14
1.4 Conclusion	1-15
SECTION II METHOD OF APPROACH	2-1
SECTION III MECHANORECEPTORS	3-1
3.1 Stretch Receptors	3-2
3.1.1 Carotid Sinus and Aortic Region	3-3
3.1.2 Atrial Stretch Receptor	3-15
3.1.3 Muscle Spindles and the Golgi Tendon Organ	3-16
3.2 Pressure and Touch	3-22
3.2.1 Pressure	3-23
3.2.2 Touch	3-26
3.3 Acceleration Receptors	3-28
3.3.1 Statocyst	3-29
3.3.2 Utriculus and Semicircular	3-31
3.3.3 Lateral Line Organs	3-33
3.3.4 Halteres of Diptera	3-34
3.4 Sound and Vibration	3-38
3.4.1 The Ear	3-38
3.4.2 Weberian Ossicles and Swimbladder	3-42



3.4.3 Lateral Line Systems	3-43
3.4.4 Lyriform Organ	3-44
3.5 Conclusion	3-45
REFERENCES	3-46
SECTION IV PHOTORECEPTORS	4-1
4.1 The Human Eye	4-9
4.1.1 The Retina	4-9
4.1.2 Optical System	4-15
4.1.3 Oculomotor Apparatus	4-15
4.1.4 The Information Content of the Retinal Image	4-21
4.2 Animal Eyes	4-25
4.2.1 The Frog	4-25
4.2.2 Arthropod Eyes	4-26
4.3 Applications	4-28
4.3.1 Photodetection	4-28
4.3.2 Oculometry	4-33
REFERENCES	4-55
SECTION V CHEMORECEPTORS	5-1
5.1 Taste	5-2
5.1.1 Detection of Stimuli and Transmission of Response	5-2
5.1.2 Mechanisms of Taste	5-3
5.1.3 Physical Property Correlations of Taste	5-13
5.1.4 Chemical Constitution and Taste	5-16
5.1.5 Sensitivity	5-20
5.2 Odor	5-22
5.2.1 Electrophysiological Considerations	5-23
5.2.2 Mechanisms of Odor	5-23
5.2.3 Absolute Sensitivity of Odor	5-31
5.3 Summary	5-34
REFERENCES	5-36



SECTION VI	THERMORECEPTORS	6-1
REFERENCES		6-4
SECTION VII	ELECTRIC AND MAGNETIC RECEPTORS	7-1
7.1	Electric Receptors	7-1
7.2	Magnetic Sensors	7-11
REFERENCES		7-12
SECTION VIII	NEURON TRANSMISSION	8-1
8.1	The Membrane Theory (Bernstein of Nervous Conduction)	8-1
8.2	Application of the Neuron Transmission System	8-4
8.3	The Neuristor	8-10
8.4	Practical Realization of a Neuristor	8-13
REFERENCES		8-16
SECTION IX	SUMMARY	9-1
SECTION X	RECOMMENDATIONS	10-1

LIST OF FIGURES

<u>Figure No.</u>		<u>Page</u>
3-1	Location of the Carotid Sinus	3-4
3-2	Arch of the Aorta	3-5
3-3	Diagrammatic Representation of Cardiovascular Reflex System	3-7
3-4	Discharge of Impulses from two End Organs	3-9
3-5 A	Response of Single End Organ in Carotid Sinus (Rabbit) Stimulated by Constant Pressure	3-10
B	Response vs Pressure Characteristic of a Single End Organ in Carotid Sinus of the Rabbit	
3-6	The Temporarily Complete Cessation of Discharge from a Carotid Sinus Receptor Following a Decrease in Pressure from 120 MM Hg to 30 MM Hg and its Subsequent Return	3-11
3-7	Diagram of an Electronic Analog of the Carotid Sinus and the Equation Which Describes its Function	3-14
3-8	Response of an Atrial Fiber to Various Levels of Filling in an Isolated Left Atrium. Impulse Frequency at Each Point was Averaged over 0.5 sec. of the Record 1 sec. after each Injection.	3-17
3-9	Rotation of Muscle Spindles and Tendon Organs to Muscle Fibers	3-20
3-10	Pacinian Corp.	3-24
3-11	The Membranous Labrinth	3-32
3-12	Diagram to Illustrate the Mechanics of the Haltere Articulations	3-35
3-13	Cross Section of Cochlear Canal	3-40
4-1	The Human Eye	4-3
4-2	Longitudinal Section of Central Macula of Human Eye	4-4
4-3	Photochemical Reactions of Rhodopsin	4-11



<u>LIST OF FIGURES (Cont.)</u>		<u>Page</u>
4-4	Distribution of Rods and Cones in the Human Retina	4-12
4-5	Foveal Acuity	4-14
4-6	Illumination of a Single Retinal Cone	4-16
4-7	Scale Drawing of Optical System of the Human Eye	4-17
4-8	Acuity as a Function of Angle from Fovea	4-18
4-9	Calculation of N	4-22
4-10	Ommatidium	4-27
4-11	Illumination at Eye	4-29
4-12	Typical Command Control System	4-38
4-13	Command Control System With Oculometer	4-40
4-14	Block Schematic and Flow Graph of Oculometer Command Control System	4-42
4-15	Oculometer Command Control System: Operators View Before and After Correction	4-43
4-16	Operators Magnified View of Target and Missile at Impact, Range 2 miles	4-44
4-17	Hi-Fi TV System	4-50
7-1	The Electric Field Around a Fish	7-3
7-2	The Electric Field in the Presence of an Object: (A) Of Low Conductivity, (B) Of High Conductivity	7-14
7-3	Threshold Field as a Function of Pulse Width and Frequency	7-9
8-1	Illustrating Nervous Conduction According to the Membrane Theory	8-3
8-2	Postulated Neuristor Recovery Curve	8-12
8-3	Neuristor Sections	8-15



1. Introduction

This report summarizes the results to date of a review of biological mechanisms for application to instrument design, performed for the National Aeronautics and Space Administration, Office of Life Science Programs, under Contract NASR-16. The report is published in two documents:

ARA-1025 Summary Report on a Review of Biological Mechanisms
for Application to Instrument Design, and

ARA-1026 Bibliography on Biosensors, a Sampling of the World
Literature 1900-1961 (Second Edition).

The compilation of ARA-1026 accounts for a considerable portion of the effort expended during this program. It is an extensive bibliography compiled through a survey of the literature covering the field of biosensors. After a substantial amount of references had been accumulated for the bibliography, an interdisciplinary analysis of the collected data was commenced for the purpose of revealing information of interest to instrumentation designers. Progress in this preliminary analysis is reported in the present document, ARA-1025.

It must be emphasized that the data analysis reported here has been a pilot program, in which several approaches have been tried and evaluated. This pilot program has provided the project team the opportunity to resolve many of the difficulties that naturally arise in a study of this sort. These difficulties primarily concern the techniques for assimilation, review, and interpretation of biologically-slanted data and their translation into terms and groupings useful to the engineer. In the relatively short time available for the analysis, a successful methodology has been evolved which can be employed in a continuing program to provide a comprehensive guide for the adaptation of biological systems to instrument design. The pilot program reported here provides examples of techniques which, it is believed, provide a useful basis for continuation of this work on a broader scale.



1.1 General Background

It is very natural - indeed inevitable - that one ask what value can derive to the benefit of instrument design from a generalized study of biosensors. It must be said frankly that there are some very capable designers that will never be able to benefit from such a study. But of this latter statement, the same may be said in connection with the transistor, the laser, and perhaps the "neuristor" described in a later section. Nevertheless, our technology requires a constant influx of new concepts and knowledge. The pressures being brought to bear on our space exploration programs are accelerating the need for advances in many areas of instrumentation. There is little assurance that all these needs will be adequately met by the conventional processes of refining or inventing strictly "physical" measuring devices and systems. What is hoped for is a new approach to instrument development that has the potential of yielding superior classes of instruments to anticipate and meet some of these new needs.

Among those concerned with the development of new instrumentation, one may find two fundamentally distinct dispositions. Many workers are disposed toward doing the best possible job within the framework of the currently recognized technology. The virtues of this approach are not to be dismissed. But it is those who are disposed toward the exploitation of radically new concepts that will profit from the approach represented by this study.

One need not become involved with theology to appreciate that nature represents a truly remarkable technology - a technology that has evolved quite independently of our own, or of our system of thought. Except for the fact that we have become conditioned to its presence and have gained some very fragmentary information about it, one might compare it to a technology developed by a newly-discovered extraterrestrial society. Now certainly, if such a society were to come to our attention, we would spare no effort to acquire the most detailed knowledge possible about what they do and how they go about it. It would be very difficult, of course, because their methods would have been conceived in non-human "minds" and with non-human "motives". Their techniques, in many cases, would simply not make sense to the human engineer. Because of these difficulties, the undertaking would be a costly one; but we would do it nevertheless. Why not,



then, undertake the relatively simpler task of exploiting that technology represented by biological organisms, specifically by biosensors? They are here, are generally not hostile, and we do have a small beginning toward their understanding.

The basic objective of this program is to provide the field of instrumentation with a new area of insight and techniques that may be exploited to advance the state of the instrumentation art. The program consists of a study in biosensors - not a philosophical comparison of the world of nature and the field of instrument engineering. However, a few such comparisons are helpful in setting the background for the biosensor study and in understanding some of its potential utility.

Perhaps the best way to commence this line of investigation is to examine some of the fundamental differences between the ways that things are done in nature and in instrument technology. It is necessary, first of all, to make a distinction between the methods employed in nature, which will be termed the biotechnology, and the study by man of those methods, which is conventionally called biology. With this distinction in mind, it is then possible to compare the biotechnology with our own technology. These may be compared conveniently on the basis of five characteristics: design, materials, fabrication techniques, control mechanisms, and evolution of the technologies. These will be taken up in order.

Design

In the human technology, design often reflects the mental processes of the designer. An engineer usually likes to build his devices, no matter how complicated, out of simple well-understood building blocks. In this way, his mind can cope with the design, fabrication, and maintenance of very complex systems. The biotechnology, on the other hand, apparently employs completely different "design" methods. While we do perceive the use of building blocks of a sort in nature, the techniques and motives in biological design remain largely a mystery. In contrast to the drawing board approach associated with human design, it is said that organisms evolve by a trial-and-error process involving random mutations and natural selection.



Materials

The differences in the materials of the two technologies are more readily described and understood than the differences in design technique. The human engineer builds his devices from uniform, macroscopic pieces of material. If a product of our technology were to be examined under a microscope capable of showing individual molecules, no pattern other than the natural crystal structure would be seen in the arrangement of individual molecules, or small groups of molecules. It is not until very large numbers of molecules are observed at one time that a humanly designed pattern will be discerned. It is to be noted that pattern, form, or detail structure - contrasted with uniformity or a random mixture - is an essential characteristic of any product of our technology and probably also of the biotechnology. (The existence of a detailed pattern in the particles of certain meteorites has, for example, been taken as an indication that they are possibly fossils of extraterrestrial life forms.) In the products of the biotechnology, pattern or structure will often be observed within the assemblages of individual molecules, small groups of molecules, or cells. In fact it seems that as larger and larger groups of molecules are observed in organisms, the less complex the patterns and structure become - which is in direct contrast to what happens in our own technology. For example, in many animals there are pumps, valves, pipes, etc., that are rather similar in broad outline to the relatively simple hydraulic components of our own technology. The fine detail or pattern of the structural tissue of these components is however, very complex, whereas the fine structure of the metals, plastics, etc., out of which pumps and valves are made in our technology is very simple. It is also to be noted that the products of our own technology are, in general, much more closely controlled than those of living things. Thus the difference between two Boeing 707's of the same type will be very small. However, there will be substantial differences between, for example, two elephants - even if they have the same parents.

It is probably an essential corollary of differences already noted that the basic materials of the two technologies are different. The materials of the biotechnology are made up of very complex molecules. They are stable only over a very narrow range of physical conditions. They exist, very often, only as a



result of the myriad processes of life and thus life support is often essential for the maintenance of these materials in their operating conditions. A particular limitation of the materials of biotechnology is the absence of good electronic conductors and insulators. In our own technology, materials are, on the whole, made up of very simple molecules. They are stable over a wide range of physical conditions and generally require nothing at all comparable with the life support involved in biological materials. Electronic conductors and insulators make possible the transmission of power and information by electronic conduction.

Fabrication Techniques

The differences already discussed between the two technologies are obviously very great. However, one extraordinary difference is in the methods and motive power of fabrication. Bio-materials are able to grow themselves by spontaneous selective grouping of molecules. This growth appears to be under the control of very complex "blueprint" molecules. The materials are fabricated and deposited in an organized way to make up the complex macro-systems of the organism. For example, in the growth of the human eye many complex materials have to be manufactured and then built up to construct the total "eye system". In complete contrast, the basic materials of the devices of our technology have to be collected, purified, machined, assembled into systems, tested - and all these processes have to be directly or indirectly under the control of a human. Consider for example the material resources and many human skills required to make a TV camera - admittedly only a very poor analog of the human eye. In the biotechnology the eye grows spontaneously. There seems to be absolutely nothing in our technology, not even of the most primitive nature, that is in any way analogous to the fabrication technique of growth which occurs in the biotechnology. The technique of crystal growing, which does involve direct fabrication of a macro structure from individual molecules, cannot be considered equivalent to biological growth, as the pattern or structure of the resulting crystal is not under our control.

In addition to the differences in material and motive power of fabrication, there is the interesting item of scale. This can be stated simply. Nature always builds up from the cellular level regardless of the physical size of the product.



Man, on the other hand, is generally required to build down, that is to begin with gross units of material which are then cut, machined, or otherwise processed to remove excess material. As the product becomes smaller, it becomes more difficult to handle and the process becomes more difficult to control.

This characteristic of the human technology has probably evolved because of the physical size of man, since his size limits the spatial resolution of his motions. His methods of fabrication have generally involved manipulation of the materials, in at least one stage of manufacture, either manually or by machines constructed to his own or a larger scale.

In recent years these limitations have been eased somewhat. Man has learned to manipulate material using smaller-scale methods such as the micro-manipulators employed in biological research. He has also learned to manipulate on the atomic and subatomic level by means of electrical and magnetic fields.

It appears that, until the secret of controlled growth is resolved, the best opportunity for construction on the microscopic level lies in these latter techniques - material manipulation by electric and magnetic fields. This idea is discussed further in Section 1.3.4.

Control Mechanisms

As in our own technology, control mechanisms play a vital role in bio-systems.

A control mechanism can be considered, in general terms, as consisting of a source of information (i.e. an input from the environment), information transmission circuits, an information processing function and an information output device (i.e. a means with which to effect the environment). Of these four items, it is the mechanism of information processing in which there is the greatest contrast between our own and the biotechnology. Thus, for example, the eye is somewhat similar to a TV camera, the optic nerve analogous to a multi-wire telephone cable. Skeletal joints and muscles are not too different from electro-mechanical systems in our own technology. However, the data processing function (e.g. pattern recognition) that is performed by the brain (and retina) is of a nature and degree completely different from the data processing performed by machines of our technology.



There is relatively little knowledge of how the brain works but it seems that unlike digital computers, it is capable of processing simultaneously large amounts of data. Digital computers operate only in a sequential manner.

It seems evident that brain action must involve a tremendous amount of parallel data processing, since it is able to perform prodigious feats in a very short time even though the response time of its neural elements is relatively very long. Although work is being done on brain and the operation of very large scale complex logic networks.

Evolution of Technology

Since we do not yet fully understand the biotechnology we perhaps cannot say for cert. in why it has developed the way it has. It will undoubtedly be of the very greatest interest to compare it with any similar occurrence on other planets. Our own technology is the way it is - not necessarily because it is the best or only technology - but because the limitations of our mental and physical processes, knowledge of the physical world, and our engineering techniques have more or less constrained its development into its present form. One could say, perhaps, that it has evolved by a natural selection of our physical theories and engineering techniques -- the selection involving the limiting parameters of our human existence.. In particular we may consider two of the most important of our inherent limitations -- our physical size and our methods of thinking.

We are macro-beings and, like our caveman ancestors, have fashioned most of our tools and devices more or less by hand, employing materials that are found lying about on the earth. The limitations imposed by our physical size, and by the range of materials generally available for fabrication, is strikingly evident when we consider how a working analog of the eye could be constructed in our technology. Possible the nearest device to it that we have now is the image orthicon TV camera tube - although, of course, it is much larger - especially with its ancillary equipment - requires far more power than the eye and lacks many of the refinements of human vision.



Secondly, our methods of thinking are very restricted. In order to function well, the brain must be able to simplify and to clarify; if necessary, by constructing simplified models of reality. Faced with large scale complex problems the brain quickly becomes "baffled" and cannot function. Many of the great scientific achievements in our history have come when men have gathered the complex threads of experience into a simple comprehensible pattern. A prime example, of course, is the work of Newton in mechanics. In engineering, it is a common, and very sound practice to work with theories, models, design techniques, etc., that may be known to be incorrect or limited in some way in these cases where a completely rigorous approach proves difficult. In this way engineers can produce something of value while waiting for the pure scientists and mathematicians to get nearer to a perfect analysis. As an example we may quote the case of filter and network design in communication engineering. For the last forty years or so, filters have been a vitally important part of large communications systems. They have been designed, however, for most of this period according to the well known image theory which is based on assumptions known to be untrue and which yields circuits that are known to be inefficient and uneconomic. Nevertheless, experience has shown that the method yields acceptable results and is very easy to understand and to apply. The general problem of getting the best possible performance from a given number of components has yet to be solved. By accepting a restrictive type of network (i. e., the ladder) modern work on network theory has shown how an optimum circuit can be designed.

The limitations of our thought processes have not unduly restricted the development of our technology up to now. Thus in the case of filter design, the difference between an image designed filter and one designed according to modern network theory is not very great. The reason is that generally a filter



is not a very complicated structure - when analyzed mathematically it usually involves only low order determinants. However when we do come to deal with inherently very complex devices - such as pattern sensors - a 'simple' approach may not be adequate, yet the fully rigorous treatment is often beyond us.

Biotechnology presents us with a chance to break this deadlock as it contains, in proliferation, very effective complex systems. It is not at all suggested that study of complex bio-systems is going to be easy; quite clearly it will not, for our mental processes are not tuned to the type of work that has to be done. However it does provide us with a change of learning at first hand the basic principles that must be involved - just as a study 50 years ago of birds could have brought us sooner to an understanding of the principles of aerodynamics.

The fact that most of our devices up to now have necessarily been macro-structures has also not been a severe limitation. However future development should certainly be towards micro-techniques: to quote but one example - the day is approaching when the speed of computers will be limited by the time it takes to transmit information over the physical dimensions of the whole computer. We can learn from the bio-world many new principles that must be understood if we are to develop a micro, or molecular, version of our own technology.

1.2 Scope of the Study

The present program comprises a review of biological mechanisms, notably biosensors, for application to instrument design. Commencing with an extensive survey of world literature, the study has broadened into an evaluation of specific sensors and classes of sensors for the purpose of deriving ideas that might be transposed into the field of instrument design.

The remainder of this introductory section is devoted to discussion of how the results of a study of biological mechanisms may be employed to improve our own instrument technology.



1.3 Exploitation of Biological Mechanisms

The exploitation of biological mechanisms may be considered in the light of four general approaches:

1. the direct imitation of biomechanisms in the materials of our technology. For example, one may seek to imitate bio-inertial sensors to assist in the further development of accelerometers, gyros, stable platforms, etc.
2. by utilizing the living biological device directly, either as part of the complete organism, or maintained separately by artificial means and incorporated into an instrument system. This would include the use of the brain of an animal, or of man, as a pattern detector.
3. by studying the basic principles of operation of biosensors to see whether the idea can be carried over into our technology. In this case the natural performance parameters of the sensors are relatively unimportant since they are a function of the technology in which they will operate; it is the idea that is to be taken. A sensor that is relatively insensitive in its natural environment may be enhanced by our technology to yield a very sensitive instrument.
4. by developing the capability for fabrication and synthesis at the molecular level, i.e., by developing a molecular technology that is basically similar to, but very much simpler than, bio-technology.

These four approaches are developed further in the following sections.

1.3.1 Direct Imitation of Biological Mechanisms

It would, of course, be premature to make any kind of a final judgment on the probability that any sensor mechanisms can be profitably imitated in the present instrument technology. Opinions on this point have diverged somewhat among the project personnel; all, however agree that the preliminary nature of the evidence gathered during the current study precludes a final judgment either way.



The probability that a particular instrument will be substantially improved by making a physical copy of its biological counterpart is quite small. The biosensor and the physical instrument are the products of two widely different technologies, as already described. The chances are that both reflect nearly the best available design in their respective technologies; but the excellence of each is only relative to its respective technology and has no ex officio status in the other. There is therefore no assurance that a particular biosensor "design" can be successfully carried over into the instrument technology, with its different materials and methods of fabrication.

It appears that this approach is less apt to be profitable than that of utilizing the basic principles of operation, discussed in a later section. This point is illustrated allegorically by Harvey E. Savely (Ref. 1):

"In discussing living systems as prototypes for engineering application ... I would be less than fair if I did not warn you of a conservative attitude that you may encounter among some biologists..."

"..... Imitation of nature's solutions may not be economical or practical. We have moved ahead in our technology by getting our hands on underlying principles and applying them in ways that may not exist in nature."

"Compare bird flight to airplanes, for example. Imitation here did not get us very far. A basic research program 60 years ago in the structure of feathers would have led us up a blind alley. It was the understanding of the physics of air flow that spurred on our development. In the same way we must look for the fundamental principles at work in the nervous system, and not be tempted to imitate what may turn out to be the 'feathers'." (Ref. 1)

1.3.2 Utilization of Living Biosensors

The development of this technique probably lies rather far in the future. The possibilities are naturally quite exciting to many biologists, and for many reasons commend themselves to the attention of engineers contemplating future developments in instrumentation, especially in those cases wherein a particular biosensor performs better than any comparable instrument, or where the instrument does not exist. There are, however, several practical problems that must be considered.



From the point of view of the instrument engineer, the biosensor (including perhaps the entire organism) must in many instances be treated as a "black box", the mechanics of which are neither thoroughly understood nor precisely controllable. This is not good engineering practice, although it is frequently tolerated in the absence of an alternative. A second problem is that of bridging the two technologies in order to devise a coupling mechanism between the biosensor and the instrument. Finally, the problem of keeping the tissue of organism sufficiently "alive" to perform its function, though not impossible, may cause considerable inconvenience. This latter problem, however, is constantly being pursued by the bioscientists for well-established motives other than those considered here.

There is an organism - the human being - that is frequently used as a "black box" in an otherwise inanimate system. For example, the human operator in an antitank guided missile system (such as the SS-10) functions in this way. In this system the operator has to perceive the spatial relationship between missile and target and then transmit appropriate control signals to the missile. In almost all other guided missile systems this task is performed by inanimate equipment. However, in the case of the antitank system no means has been found of reliably identifying a tank on the battlefield. The human is able to do it by means of his pattern recognition capabilities, and it is for this reason that he is included in the system. This pattern recognition faculty is a brain function, and thus, from a pragmatic point of view, the brain itself may be considered as a detector - a detector of pattern.

An important aspect of the use of a human pattern detector is the means employed to couple him to the rest of the system. It is not presently practicable to make a direct physical connection into the neural circuits of the body, so that resort has to be made to neuro-muscular/mechanical coupling, such as a hand-held control stick. The oculometer system described in Section 4 does allow, in effect, a connection to be made to the neural system of the anti-tank-missile-operator, just at that point in his perceptive system where the position of the "pattern" (i.e. the tank) is made known to his consciousness.



Life support systems for manned spacecraft offer another wide field of possibilities for direct application of living sensors. By monitoring the reactions of living biosensors or complete organisms, one may be able to determine carbon dioxide or ozone content in a atmosphere, toxic accumulations in atmosphere or in closed-cycle water systems, radiation levels, as well as many other crucial parameters of the new environment.

1.3.3 Utilization of Basic Operational Principles of Biomechanisms

This approach may provide the earliest benefits from exploitation of the biotechnology. The general technical difficulties associated with the utilization of living biosensors do not apply here.

It can be argued that our present instrument technology is so effective, because of the pressures that have kept it creative, that it is not apt to benefit from any new ideas that may accrue from a study of the biotechnology. Many of our instruments have been in existence for some time now and have evolved to a high degree of performance and reliability.

One of the first objections that can be raised against this point of view is the very obvious one that no field, least of all instrumentation, is static. It is always being presented with requirements of increasing severity and there is no assurance that present approaches will meet these needs adequately. For the first time, the dramatically different environment of space will soon be ours to explore, and will thus open up an entirely new spectrum of instrumentation requirements. In addition to the much broader varieties and ranges of variables to be measured, there is the increased difficulty of placing and maintaining the instrument in an operating position. Certainly there has never been a time when radically new ideas were needed as much as they will in the very near future.

Even in the familiar earth environment, there are vast areas for exploration that are certain to require major advances. Just two examples are apparent in polar and undersea expeditions.

Finally, the frontiers of data handling and processing provide a host of new requirements that, it is already widely acknowledged, will require more than conventional approaches to meet. We simply do not understand the basic operating principles of many of the machines that we already want to build - for example,



machines to read handwriting - and if we did understand them we could probably not build them. The study of pattern detectors in the biological world is obviously a very effective way of learning about the fundamentals involved.

Consider some of the dramatic implications of the ability to exploit bistable (on-off) phenomena at the cellular or molecular level. If individual cells or molecules could be made to exhibit this behavior, and if a way could be found to monitor their individual states, then logical elements built on this basis could be used to construct high-density complex networks. These could be used both in computer and instrumentation technology, and in the study of highly-developed logical processes.

1.3.4 The Development of a New Technology

The two main stumbling blocks to direct exploitation of the biotechnology are the differences in materials and in fabrication techniques. If these differences could be narrowed, the way would be open for more fruitful exploitation. In a sense the natural progress of bioscience - in which progressively better understanding of all life processes is being obtained - is helping to narrow the gap.

Biotechnology is, in essence, a molecular technology, since many of the processes and systems involve either small groups of molecules or else have a significant pattern or, structure at the molecular level. Therefore another approach to exploring biotechnology would be to develop our own form of molecular technology.

The first stage in such a development must be to develop techniques of fabrication. Because we are macro-beings, almost everything we can make is a macro-structure. New techniques will be needed to enable us to build organized micro-structures. In the biotechnology, such fabrication is done by growth. We simply do not understand how this happens, and until a breakthrough comes we must use other methods of assembling groups of molecules in previously designed patterns and structures. Some of the techniques now being used in the semiconductor industry are tending towards molecular fabrication.

A method that might be developed to meet very closely the ideal of handling small groups of molecules or individual molecules is the mass spectrograph. In this, a beam of one particular type of ion can be selected and then magnetically or electrostatically directed to specific points on a target. Patterns or structure



could then be built up by suitable modulation of the ion selecting and ion deflecting circuits.

The development of a simple molecular technology should lead to a more fruitful cross fertilization between the physical and biosciences. In the same way that we can expect a carryover of techniques and principles from biotechnology into the simpler but analogous molecular technology, knowledge and techniques built up in this molecular technology should lead to a better understanding of biotechnology. By attempting to work in a very simple analog of the biotechnology, we can derive the great benefits of having exact knowledge and control over our experiments. The results will be relatively easy to interpret and our minds can fashion theories to suggest further progress. This orderly progression from simple controlled experiments, which are capable of being easily analyzed by our mathematics, to work that is superficially complex is characteristic of the way physical science has progressed.

From a directly practical point of view, the establishment of an effective molecular technology, capable of exploiting bio devices and also of functioning independently, would be of great value. The products of this new technology would be smaller and lighter, by several orders of magnitude, than analogs in our existing technology. The techniques of fabrication are almost bound to involve very much less human participation, and be capable of close control, thus the products will probably be cheaper and have very high and predictable reliability. The sensitivity of a number of biosensors is substantially beyond that of comparable devices in our own technology. It is to be expected that the devices made in a molecular technology will be able to approach, match, or even exceed the sensitivity of similar bio devices.

1.4 Conclusion

The world of living things poses a real challenge to our engineering skills. In many ways our devices are inefficient and clumsy compared to their biological counterparts. There seems to be no reason why we should not study and exploit the products of the biotechnology just as we would exploit the technology of any extraterrestrial beings that might be discovered.



However, the difficulties involved must not be overlooked. An immediate flood of simple and obvious applications from the bio to our own technology is not to be expected - for reasons discussed earlier in this section. One must keep in mind nonetheless that - understand it or not, exploit it or not - the biotechnology exists. Its accomplishments are impressive and a challenge to our openmindedness and adaptability.



2. Method of Approach

The nature of this study has been exploratory in two senses. First of all, the idea of applying biomechanisms to instrument design, although not a new one, is not a well established practice. The study must also be exploratory in its methodology, involving as it does the establishment of a line of communication among several disciplines with widely varying points of view and motivation. It was decided therefore to approach the analysis of the data in a somewhat experimental fashion. The benefits of the experimental approach are twofold. It permits the individual differences among the investigators to be matched to the characteristics of the material being analyzed and thus relieves somewhat the problem of coupling the investigator to an unfamiliar field. This serves the interest of efficiency. The second benefit is the opportunity to examine the relative efficacy of different approaches to the problem. Such an evaluation would be useful in continuing studies of this nature.

The emphasis was not particularly on deriving new data, but rather on the careful organization of existing data into a form useful for its application at a later time. This does not mean that the generation of new ideas was to be discouraged, but rather that the primary objective of this pilot phase of the study program was to be assimilation and meaningful organization of background material.

Two basically different methods were employed. For the sake of convenience we might term these the "cataloging" approach and "inspirational" approach. The first, which gives the broadest coverage to all of the available material, consists of a methodical scanning of the literature to extract any quantitative data that may appear meaningful to the engineer. This method, when carried to completion, would require that every biosensor be examined and reported upon even though its potential applicability to instrumentation might not seem immediately to warrant complete transposition of its characteristics into engineering terms. In following the second, or inspirational, approach one would simply expose himself in a more or less methodical fashion to the available literature. He would follow his own particular interests until, through association of ideas, etc., he



might perceive some particular application or class of applications of a biosensor to instrumentation. He would then develop this idea in whatever way might seem appropriate. Thus the individual employing the latter method would make use of the material from the literature search only insofar as it served his purposes in triggering and developing a particular idea, and supporting it with quantitative data. In any given instance this approach would lead either to a dead end or to a useful idea. The penalty, of course, is that substantial portions of the available literature might go unused during early phases of the study. But, after all, the real point of a study such as this is the hope that some unconventional and perhaps exotic ideas for instrumentation may result-- it is not intended to be simply an exercise in sorting and cataloging per se.

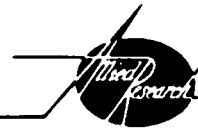
The differences between these two approaches may be illustrated by a simple analogy. Suppose that one is approaching the coastline of an unknown continent. The question would naturally arise as to the best way of exploring this wilderness. Having established an initial beachhead, one would be presented with two alternatives. The expedition could fan out, making an orderly examination of every square mile within an expanding zone. Or, alternatively, the expedition could move along at a faster pace until it came across feature of interest, such as a river. It might then abandon exploration of the coastline and concentrate on following the river to some logical stopping place, perhaps a lake. The party could then select some other particular goal of interest, such as a mountain on the horizon, and proceed toward that. The latter approach, although leaving large areas near the coastline unexplored, would be apt to acquire interesting and useful information much more rapidly. The party that elected to explore the coastline and inland zone in a grid fashion would have acquired detailed knowledge of the relatively limited area in the vicinity of the coast line. Eventually, it might expect to include in its survey those points of interest previously encountered by the more adventurous party. Thus the advantages of a thorough and comprehensive survey must be balanced against the usefulness of making some of the more exciting discoveries early in the game.



It was decided to attempt to achieve a balance between these two points of view in the present study. Although the distinctions were not severely drawn, there was a general trend toward the cataloging approach in the section on the mechanoreceptors and chemoreceptors, while the more free-wheeling technique was applied in the case of photoreceptors and neural mechanisms. The depth of coverage in this pilot study is not sufficient to permit a fair evaluation of the two methods; it appears, however, that the choice is partially dependent on the individual habits of the investigators. We are inclined to believe at this point that neither technique should be abandoned entirely, but that at least one person in a study group be given the freedom to scout new territory as the occasions present themselves.

Organization of the project was necessarily centered around the basic problem of establishing communication between the biological sciences and the engineering specialties, and in providing the investigators representing these areas with sufficient vocabulary and a basic understanding of each other's fields, to establish a workable interface. In addition to the necessity of virtually learning an unfamiliar language, each project member was faced with a requirement to put aside whatever professional prejudices he may have held toward the other field. In the case of the present group, fortunately, there seemed to be a mutual inclination to remove these barriers. Nevertheless, a predisposition toward this kind of cooperation, although necessary, is not sufficient to ensure an adequate interface. One does not acquire a sufficient insight into the areas beyond the boundary simply by agreeing to be tolerant about it. The problem of cross-education was met by persistent exposure to the material at hand, and by seeking to evolve a "middle language" through which all members of the project could communicate.

The major share of the project effort went into preparation of the bibliography, consideration of the philosophy of the approach, and orientation of the project members to what was to many of them a new discipline. The remainder of the activities were directed toward the preliminary analysis of the data, the results of which are reported in the sections that follow. It must be reiterated that in no case does the reported material comprise more than a preliminary foray into the respective subject matter, but serves simply as a sample of what approaches may be taken and the kind of data that may result from each.



3. Mechanoreceptors

It is convenient to group the mechanoreceptors according to sensed parameter provided it is realized that distinctions between groups must in some cases be made rather arbitrarily. For example, certain pressure sensors actually operate by sensing the stretch occasioned by application of pressure to the organ. Thus the pressoreceptor may be considered as a specialized stretch receptor. In some cases, of course, the details of the sensing mechanism remain a subject of controversy. The groupings employed here are informal and were selected for convenience in this phase of the study. Continuing work, in which the subject were expanded in both depth and breadth, might certainly demand a more formal grouping.

For present purposes, then, the mechanoreceptors have been grouped into the following areas:

- stretch (strain)
- pressure and touch
- acceleration
- sound and vibration

The discussions of each mechanoreceptor group will follow a format consisting of a list of biosensors in that group, descriptions of individual sensors, and in some cases quantitative data on operating characteristics. In point of the latter, it is by no means surprising that data taken in a biological laboratory are frequently expressed in terms not directly applicable to the field of instrumentation design. This fact primarily reflects the differences in approach and objective of the biological scientist and the engineer. The reduction in these differences in approach has already begun to appear as a product of the evaluation of the bio-engineer, who will undertake experiments of a biological character to answer questions of an engineering nature. (He will, of course, reverse this procedure in attempting to answer biological questions by recourse to engineering analogs and analytical techniques).

Another difficulty arises through the fact that an organic receptor does not lend itself easily to the sort of measurement and testing that an engineer likes to perform on a component in an inorganic system. Because the biosensor is so



sensitive to its operating environment such tests must be carefully tailored in terms of the situation in which it is to operate. Since most tests reported to date are slanted toward determining behavior of the sensor in its natural state, i.e., within a living organism, any quantitative data from these experiments must be considered with reservations whenever a specific sensor is contemplated in terms of its usefulness in instrument design.

3.1 Stretch Receptors

As previously noted, the grouping of the bioreceptors is somewhat arbitrary as a matter of convenience. For the present study the following were included in the stretch (strain) category.

Carotid Sinus
Aortic Sensors
Atrial Stretch Receptor
Muscle Spindle
Golgi Tendon Organ

The first two items in the list offer an example of sensors that one might be inclined to place in another category. The carotid sinus acts as a pressure sensor, while the atrial receptor senses the degree of filling (relative volume of contents) of a cavity in the heart. Nevertheless, those organs actually operate by sensing the degree of stretch of the arterial and atrial walls and hence are frequently classes as stretch receptors.

It must be recognized, however, that in much of the literature the carotid sinus receptors are referred to variously as baroreceptors, pressoreceptors, etc.

The remainder of this section presents a sampling of some of the data derived in the literature search. Several of the receptors are described and samples of available quantitative data are given.



3.1.1 Carotid Sinus and Aortic Region

Two important sets of stretch receptors are found in the carotid sinus and in the aortic area. These sensors operate, in conjunction with the elasticity of the arterial walls that contain them, as hydraulic pressure transducers. These sensors are of variable sensitivity and are crucial elements in the reflex homeostasis of blood pressure and, to a lesser degree, of respiration.

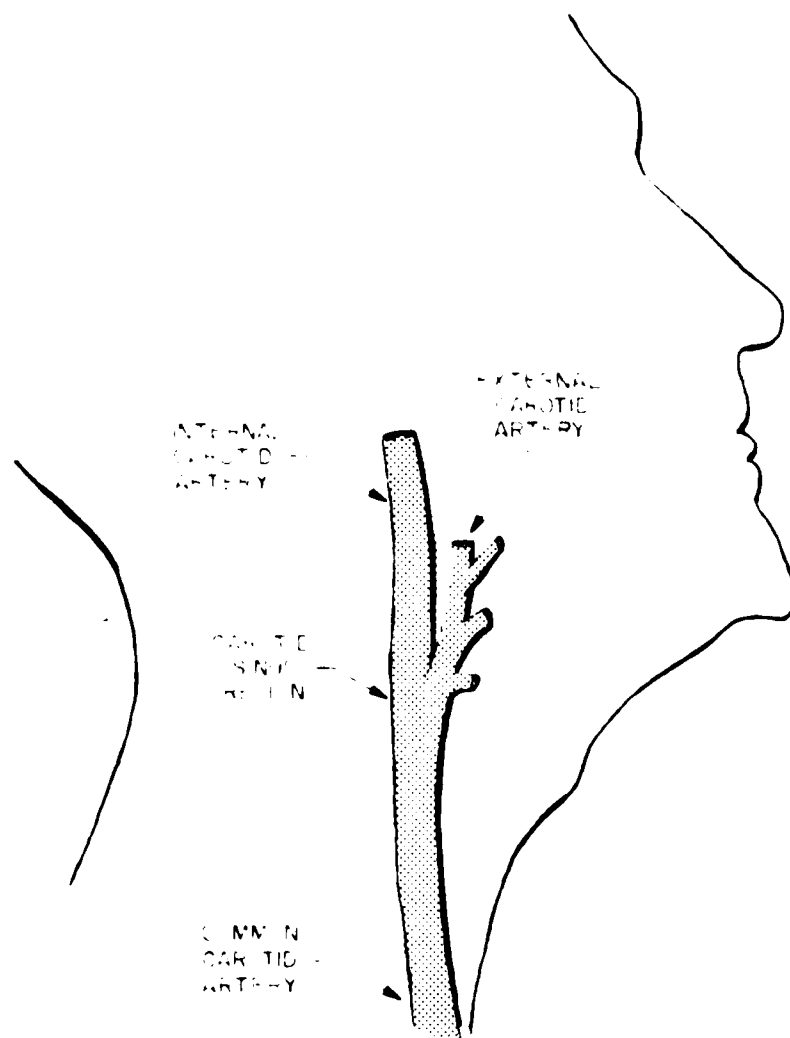
The carotid sinus is the term applied to the slight enlargement that occurs at the bifurcation of the common carotid artery into the external and internal carotid arteries (Figure 3-1).

The aorta is the main trunk of a series of vessels that carry oxygenated blood to the various tissues of the body for their nutrition. It leaves the heart at the upper part of the left ventricle and arches over the heart, descending then to the abdominal cavity (Figure 3-2). These brief descriptions of the carotid sinus and aortic region are adapted from Reference 1 in which detailed descriptions may be found.

The carotid and aortic areas both contain two types of proprioceptors responding to mechanical (pressure) and chemical stimulation, respectively. The pressoreceptors are situated among the collagenous fibers in the walls of the carotid sinus and aortic arch. They are stimulated by a stretching force, as by a rise in arterial blood pressure (Ref. 2). The afferent nerve impulses from these pressoreceptors are applied to the cardiac-inhibitory and vasomotor centers. Compression of the carotid sinus area (so as to raise the intrasinus pressure) causes slowing of the heart rate and dilation of the blood vessels, both effects resulting in a fall in systemic blood pressure. Pressure on the carotid artery some distance below the sinus (so as to reduce the pressure within the sinus) causes cardiac acceleration, vasoconstriction, and a rise in arterial pressure. These effects are accompanied by the liberation of adrenaline. The sinus and aortic nerves constitute a most important mechanism in the control of arterial blood pressure and in the maintenance of circulation to the brain. For example, these nerves are apparently the agent for compensation of the blood pressure when the body changes from the recumbent to the sitting position or from sitting to standing.

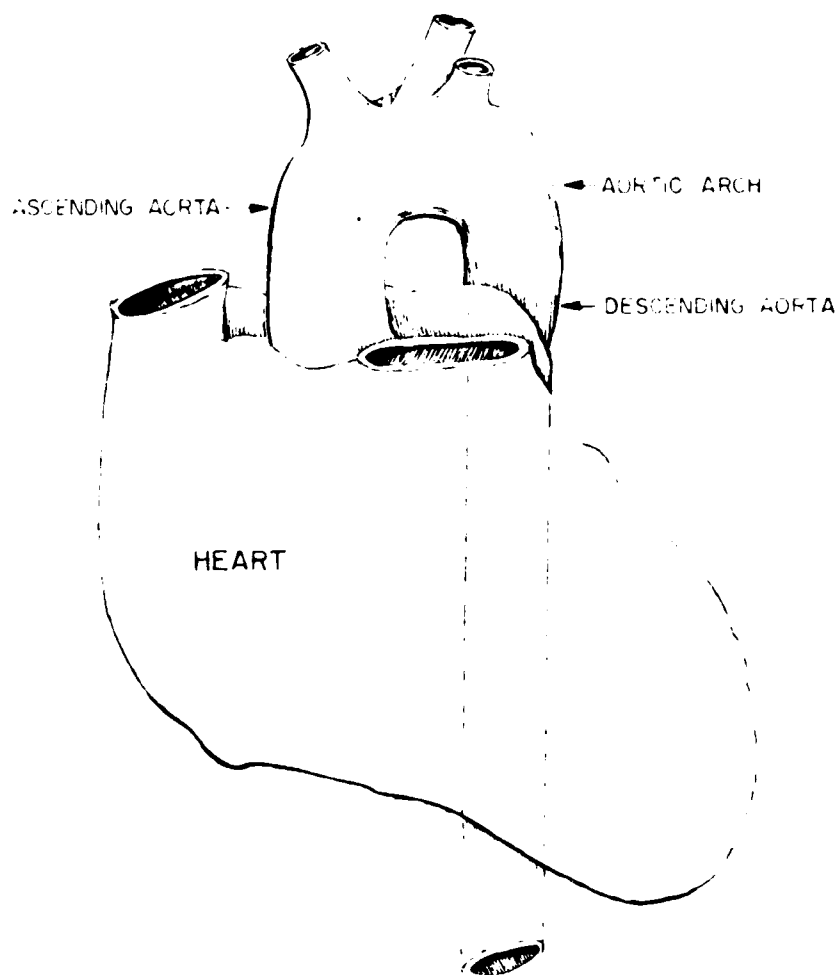


LOCATION OF THE CAROTID SINUS (REF. 1)





ARCH OF THE AORTA (REF. 1)





These mechanisms are also called into play when the blood system becomes under-filled as in the case of hemorrhage or shock, initiating a general vasoconstriction to adjust vascular capacity to the reduced blood volume and thus maintain the blood pressure. The function relationships of the sinus and aortal pressure sensors are illustrated in Figure 3-3.

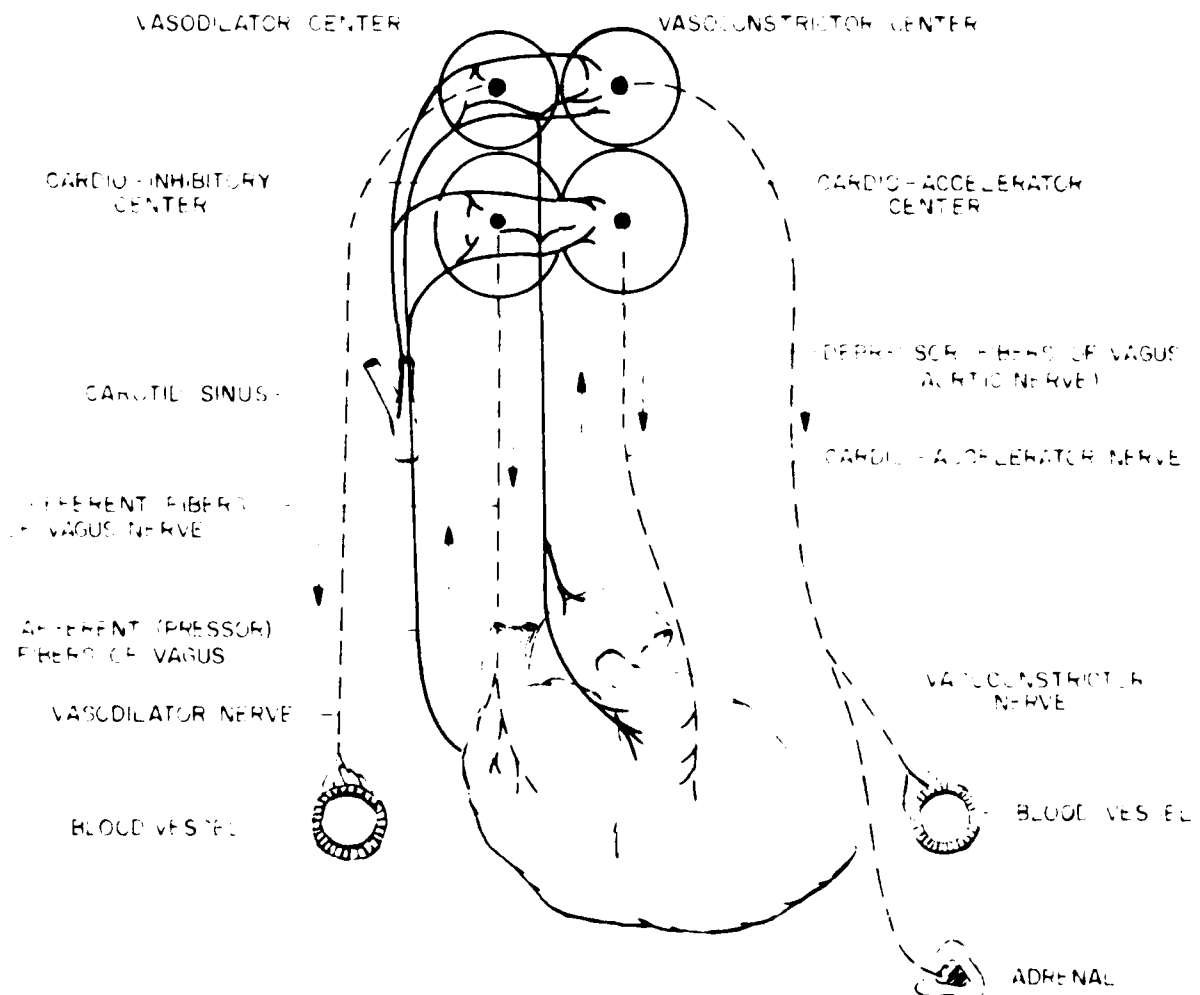
Both the pressoreceptors and chemoreceptors of the carotid and aortic areas are connected to the respiratory center as well as to the cardiovascular centers. However, it appears to the authors of Reference 2 that the effects of these sensors on respiration under physiological conditions are not generally significant, the pressure function being perhaps an obsolete mechanism no longer needed by the mammal and the chemoreceptor acting as a last-ditch defense against respiratory failure.

The arterial baroreceptors are of variable sensitivity so that the cardiovascular system is somewhat self adaptive. Experimental results indicate that the state of contraction and resistance to stretch of the barosensitive arterial walls are controlled by the sympathetic nervous system. It is thought that efferent pulses from the sympathetic nerves cause local release of chemicals that modulate the sensitivity of the pressoreceptors to intra-arterial pressure (Reference 3). This is one example of the factors that make physiological data difficult to translate into instrument data. The development of instrumentation from a living arterial baroreceptor would require specialized experiments under which all parameters were controlled to simulate the proposed operating environment. On the other hand, availability of a variably-sensitive transducer may offer some attractive advantages.

The carotid sinus contains many individual pressoreceptor end organs which differ from one another both in their sensitivity (impulse frequency vs arterial pressure) and in their threshold levels. Each responds to an increase in stretch (once the threshold has been passed) by an increase in the frequency of its output nerve impulses. Thus a rise in pressure increases the number of afferent pulses from the sinus by increasing both the frequency of discharge from individual end organs and the number of participating end organs (Ref. 4).



DIAGRAMMATIC REPRESENTATION OF CARDIOVASCULAR REFLEX MECHANISM (REF. 2)





This effect may be observed in Figure 3-4, drawn from an electroneurogram shown in Reference 4. Two end organs in the carotid sinus of a rabbit were monitored by an oscilloscope. The sinus was subjected to a pressure of 70 mm Hg, which caused one end organ to discharge at 15 impulse/sec (trace A). When the pressure was raised to 100 mm Hg, the first sensor increased its discharge frequency to 25 impulse/sec and the second sensor commenced discharge at 20 impulse/sec, the latter being shown as the larger spikes in trace B. Thus the difference in both threshold level and sensitivity of the two individual end organs are demonstrated.

The response of a single end organ to a sustained stimulus is shown in Figure 3-5A. When the pressure is brought to 65 mm Hg, the response consists of a train of pulses occurring with a high degree of regularity. The frequency of discharge falls off very slowly over some seconds and then remains steady for an indefinite time at a rate only slightly less than the initial maximum value. Compared to other receptors in general, the baroreceptors of the carotid sinus are very slowly adapting. Such behavior is quite appropriate for the regulation of blood pressure in a system wherein pulsations tend to be smoothed out at higher mean systemic pressures (Ref. 4).

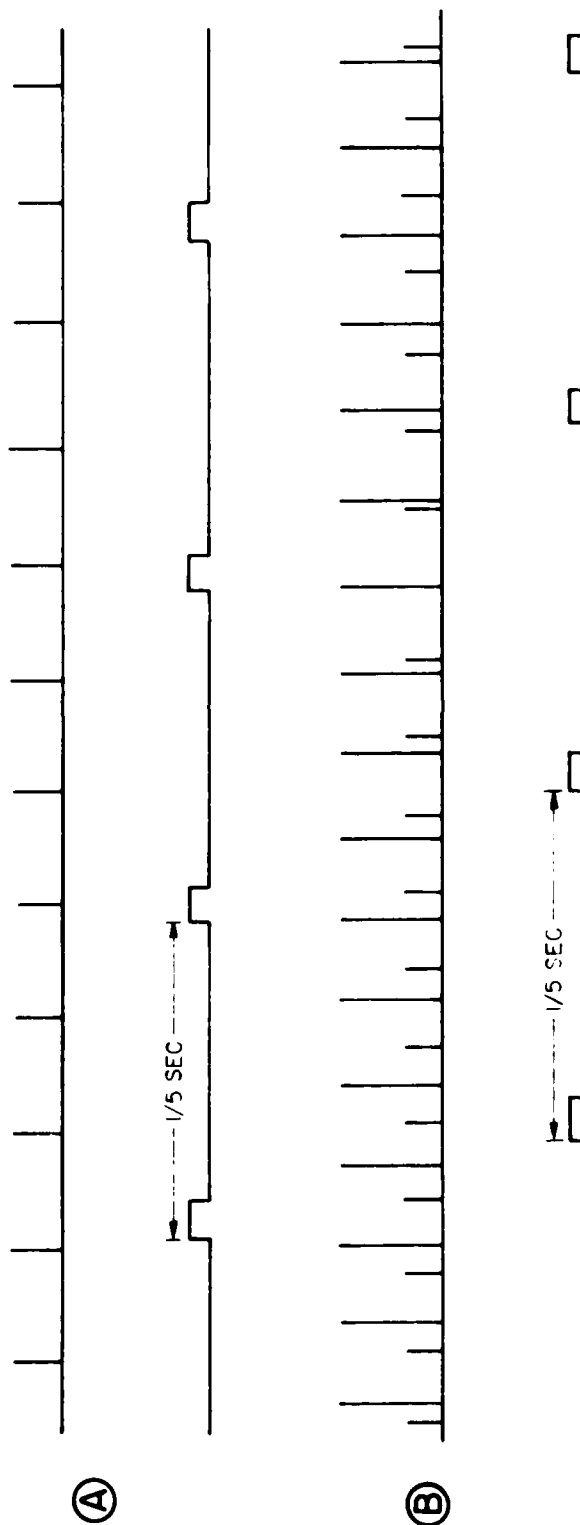
Response data in a form interesting to instrument designers is shown in Figure 3-5B. However, it must be emphasized that this curve represents only one particular end organ, and that there are wide differences in the sensitivities and thresholds of individual sensors. The curve was plotted from electroneurograms of a single receptor at four steady-state pressures. A 2-1/2 fold increase in output frequency occurred over the 5-1 pressure range. Although such a striking linear behavior over the entire range of pressures investigated is frequently observed, in other cases the sensor saturates at a level of 140 to 180 mm Hg with the maximum sensitivity occurring between limits within which the carotid sinus reflex might be expected mainly to function (Ref. 4).

The carotid sinus receptor, like many sensors, exhibits a silent period following a sudden decrease in stimulus. This effect is illustrated in Figure 3-6. When the pressure within the sinus was quickly reduced from 120 mm Hg, to 30 mm the impulse discharge stopped abruptly and completely for about 18 seconds, and then gradually developed to a new level characteristic of the lower pressure (Ref. 4). A similar effect of shorter duration is found in the case of muscle spindles.



① WHEN STIMULATED BY A PRESSURE OF 70 mm Hg.

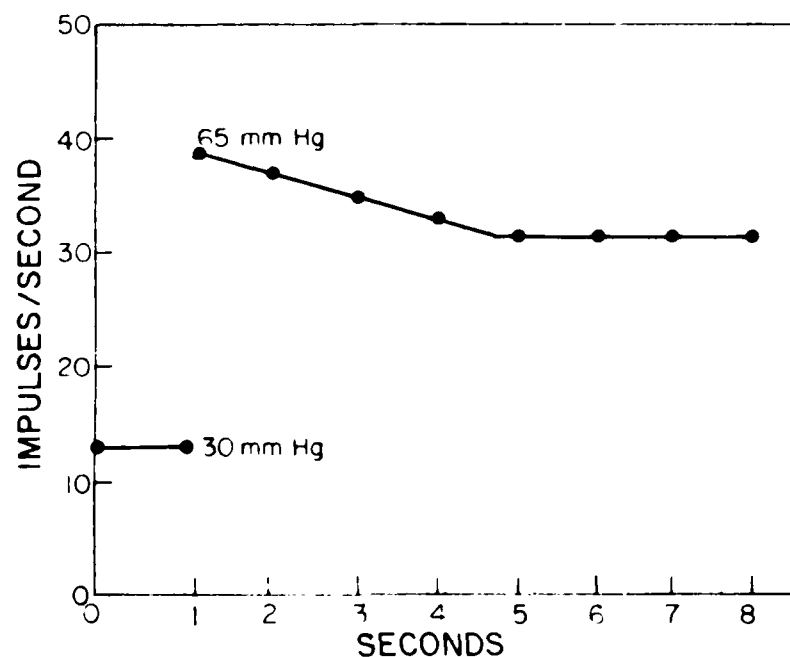
② WHEN STIMULATED AT 100 mm Hg. THE SECOND END ORGAN (LARGE PULSE) COMES INTO ACTION AT B.



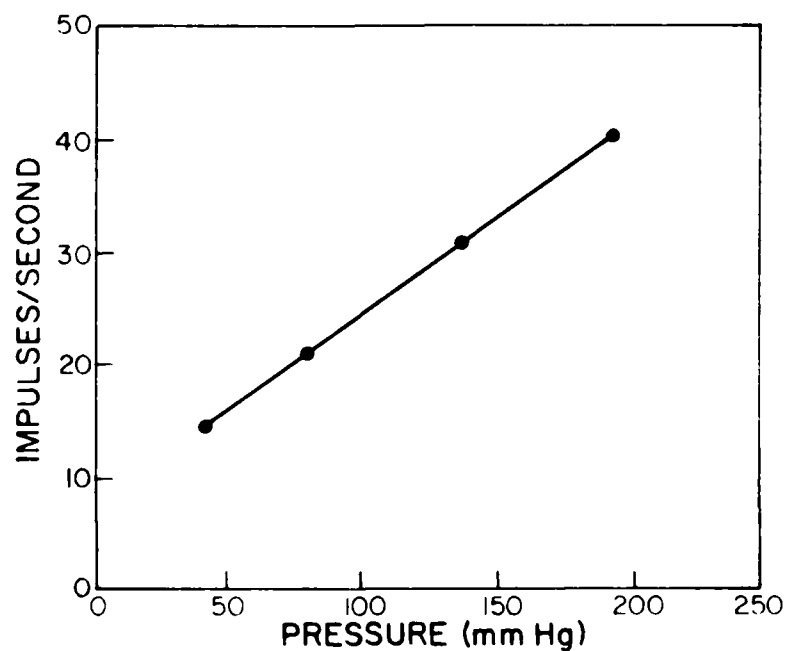
DISCHARGE OF IMPULSES FROM TWO END ORGANS
(REF. 4)



A. RESPONSE OF SINGLE END ORGAN IN CAROTID SINUS (RABBIT) STIMULATED BY CONSTANT PRESSURE OF 65 mm Hg. (FROM REF. 4)

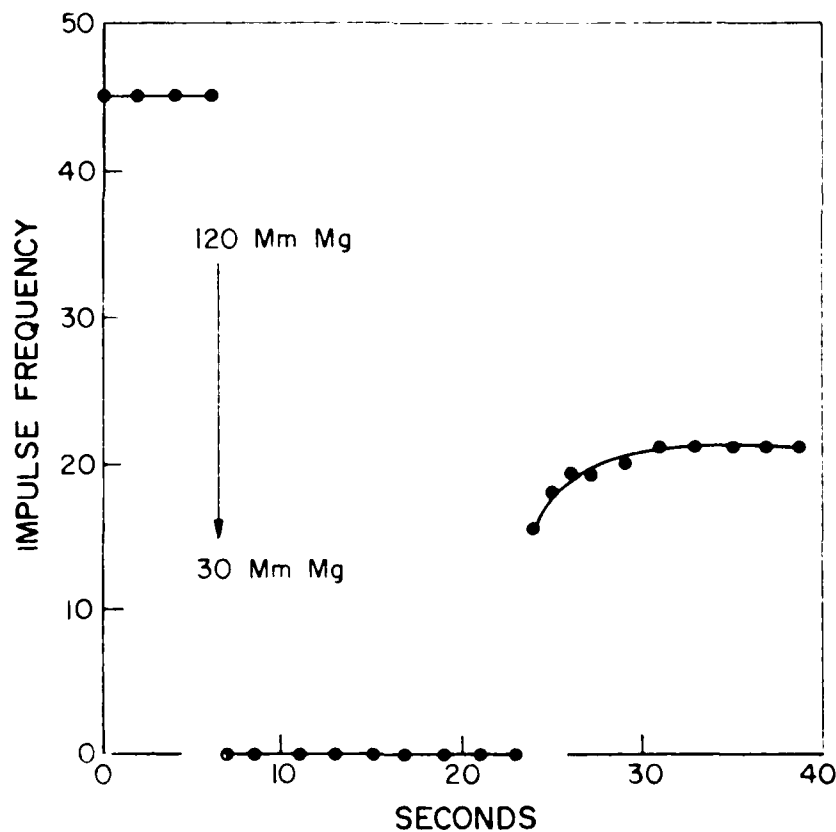


B. RESPONSE vs. PRESSURE CHARACTERISTIC OF A SINGLE END ORGAN IN CAROTID SINUS OF THE RABBIT (REF. 4)





THE TEMPORARILY COMPLETE CESSATION OF
DISCHARGE FROM A CAROTID SINUS RECEPTOR
FOLLOWING A DECREASE IN PRESSURE FROM
120 mm Mg TO 30 mm Mg AND ITS SUBSEQUENT
RETURN (REF. 4)





The quenching of sensor output by a reduction of stimulus probably contributes to the dynamic response of the organ to a pulsating stimulus. When the sinus is subjected to the normal stimulus of the arterial pulse, it is observed that at low or medium pressures the nerve impulses are produced only during systole even though the diastolic pressure is sufficiently high to evoke a discharge if maintained continuously.

The difference in carotid response to pulsatile and steady flow is demonstrated in experiments by Ead et. al. (Ref. 5). One carotid sinus of a cat was isolated and perfused by a pump in such a way that the intrasinus pressure could be controlled. The carotid sinus nerve was left intact so that afferent pulses from the sinus operated on the vasomotor center. Thus sinus stimulation acted to reduce the systemic blood pressure (hypotension) of the anesthetized animal. The blood pressure was recorded from the femoral artery while the sinus perfusion pressure was varied and the degree of blood-pressure depression was thus used as an indicator of sinus response. The results are summarized in Table 3.1, taken from Reference 5.

TABLE 3.1

Mean sinus perfusion pressure (mm Hg)	Mean systemic blood pressure (during non- pulsating sinus perfusion) (mm Hg)	Mean systemic blood pressure (during pulsa- tile sinus perfusion) (mm Hg)
0	160	160
15	160	130
80	160	110
135	115	88
240	83	83

At any given mean sinus perfusion pressure below 240 mm Hg., pulsatile perfusion is more effective in causing reflex hypotension than is non-pulsatile perfusion. At high sinus perfusion pressures (about 250 mm Hg) the two types of perfusion produced almost identical reflex responses.



An another experiment in the series, in which the natural blood flow through the sinus could be made either pulsing or steady by means of an impulse absorber, the electroneurograms for both flow types were compared. It was observed that during steady flow the impulse discharge occurs steadily throughout the cardiac cycle, while during pulsatile flow the impulse discharge is grouped, the bursts occurring during systole and early diastole of each cardiac cycle. The maximum frequency attained during pulsatile flow exceeds that seen during non pulsatile flow, although the average frequencies (or impulses per cardiac cycle) were approximately the same for both types of flow when the mean pressure was in the range of 100-160 mm Hg. It appears that a given number of pulses occurring in a short high-frequency burst are more capable of affecting the vasomotor center than the same number occurring as a steady discharge of lower frequency (Ref. 5).

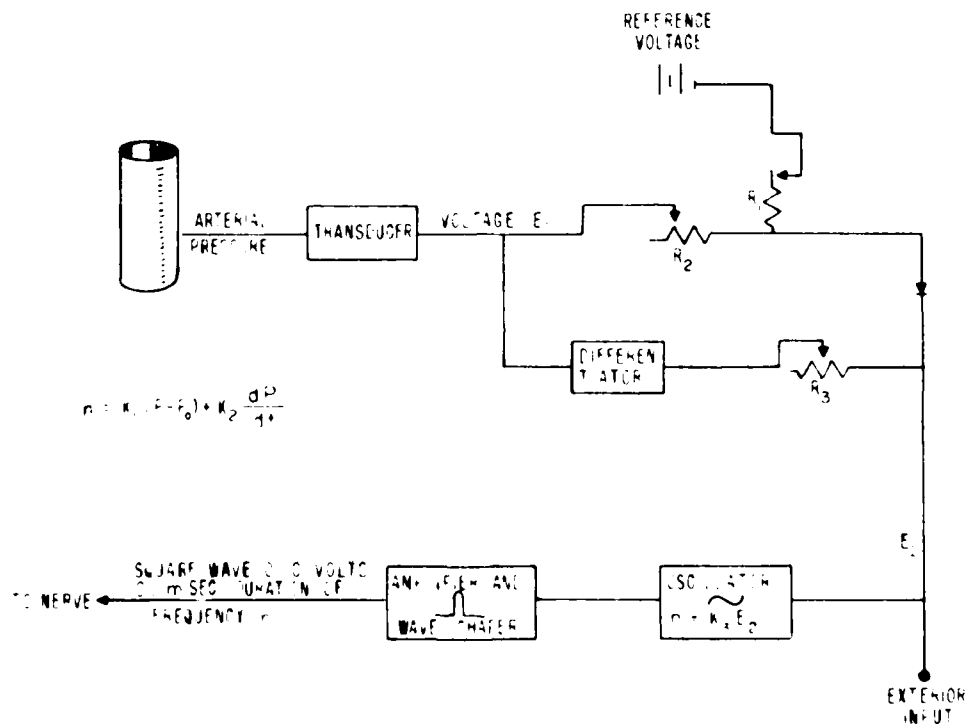
The authors conclude (Ref. 5) that the activity of the baroreceptors in the aortic area may safely be assumed to react to pulsatile changes in a manner similar to that of the carotid sinus described above.

An interesting dynamic study of the carotid sinus operating in the blood-pressure regulation loop has been performed by Warner (Ref. 6). He devised an electric analog of the carotid sinus, shown in Figure 3-7, which he substituted for the sinus. By varying the transfer function of the analog and the frequency of the pulsating stimulus he was able to measure the dynamic response of the tonic reflex system. It was observed that at certain pulsatile frequencies the effect of amplifying the effect of carotid sinus activity was to aggravate rather than diminish arterial pressure variations. Theoretically, a delay in the servo loop equal to one half the period of disturbance will change normally negative feedback to positive. The experimental results support the concept that the phase shift in the regulatory system is determined by the time lag in the response of arteriolar smooth muscle to stimulation of the carotid sinus nerve.

If at these critical frequencies, the gain around the loop is greater than one, the system is unstable and oscillations will grow. Failure to observe this instability, even when the simulator was made far more sensitive than the carotid sinus, is attributed to the nonlinear depressor response of the animal to increasing frequency of carotid sinus nerve stimulation (Ref. 6).



DIAGRAM OF AN ELECTRONIC ANALOG OF THE CAROTID SINUS AND THE EQUATION WHICH DESCRIBES ITS FUNCTION (REF 6)





From the point of view of this report, the interesting feature of the analog is the presence of a first time-derivative term in the transfer function. This term reflects the already-described properties of sensor adaptation to a sustained stimulus and the abrupt period of silence following reduction of the stimulus.

3.1.2 Atrial Stretch Receptor

The upper chambers of the heart are known respectively as the right and left atrium. These are the input areas of the heart -- the right atrium receiving venous blood from the body's circulatory system, and the left atrium receiving newly oxygenated blood from the lungs.

The exact innervation of the atrial regions apparently remains controversial. However, the nature of the receptors in these areas and their related reflex functions have been studied by many authors (References 7 through 13). For purposes of the present report, the following composite picture may be presumed.

The function of the atrial receptors appears to be twofold: (1) to adjust the cardiac rate to accommodate the rate of flow of blood into the atria, and (2) to regulate the total volume of actively-circulating blood through reflex mechanisms affecting the kidneys and other organs. An example of the first function is seen in the Bainbridge reflex, named after its discoverer. The reflex causes an acceleration in heart rate when the pressure of blood entering the auricle is raised. The reflex is carried out presumably through the afferent vagal termination beneath the endocardium and in the walls of the great veins near their entrance to the auricle. The nerve fibres are stimulated by the increased venous pressure which rises only after the cardiac chambers have been completely filled. Through this mechanism, the heart rate is adjusted automatically to the volume of blood poured into its chambers (the venous inflow) (Ref. 2).

The second function is manifest in the homeostatic response of the kidney to changes in the volume of blood circulating in the system. The receptors of the atria, and possibly the proximal portions of the veins entering the atria, are believed to react to circulating blood volume in such a way as to control the flow of urine and thus stabilize blood volume. It has been found that effects tending to increase volume, e.g. infusion, lead to diuresis (abnormal urine flow) while hemorrhage will cause oligurea (antidiuresis).



The sensors involved are variously termed stretch receptors and pressure receptors. There are indications that both may be present. However, the pressoreceptors are described as similar to those in the carotid and aortic areas (Section 3.1.1) which may be considered ultimately as stretch receptors.

The stretch receptors in the atrium have been found responsive to degree of filling of the chamber (Ref. 7). A sample response curve taken from the reference is shown in Figure 3-8.

3.1.3 Muscle Spindles and the Golgi Tendon Organs

The receptors situated in skeletal muscles and in the tendons and joints furnish information to the central nervous concerning the movements and positions of the limbs and other parts. Afferent fibers carrying this information make up from 1/3 to 1/2 of the fibers in a so-called motor nerve. As a result of the messages received by the nervous centers, the contractions of individual muscles and groups of muscles are coordinated to produce smooth, finely adjusted and effective movements which would be impossible in the absence of such guidance from the periphery. For this reason the term kinesthetic is applied to this group of receptors. A proportion of these afferent impulses arouse no sensation, their information being delivered to centers lying beneath consciousness. To others are due the sensations grouped under deep sensibility. The receptors in the situations mentioned respond to mechanical stimulation, e.g., pressure or stretch. These types of stimulus are furnished by the strains and stresses set up in the muscles, tendons and joints during muscular contraction (Ref. 2).

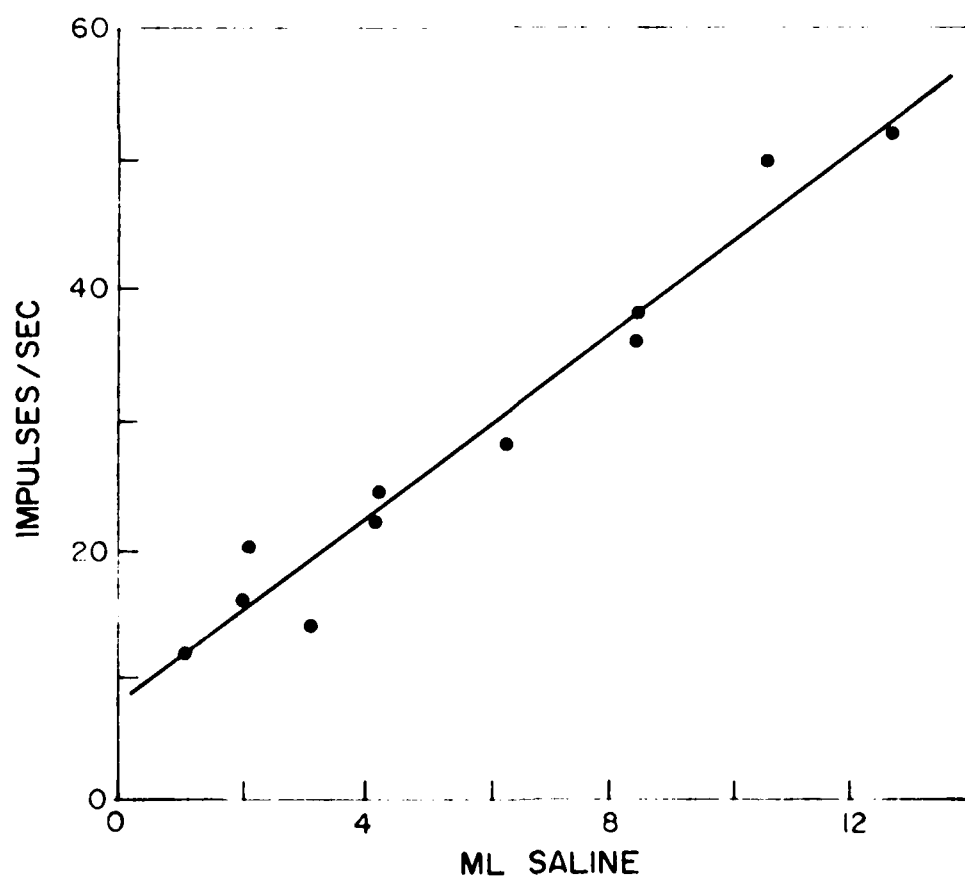
The sensory endings in the various situations mentioned above are of four main types:

- muscle spindles
- Golgi corpuscles
- Pacinian corpuscles
- free nerve endings

The first two of these are considered to be stretch receptors, and as such will be discussed here. The third is a pressure receptor, discussed elsewhere, and the fourth is a pain receptor.



RESPONSE OF AN ATRIAL FIBER TO VARIOUS LEVELS
OF FILLING IN AN ISOLATED LEFT ATRIUM. IMPULSE
FREQUENCY AT EACH POINT WAS AVERAGED OVER 0.5
SEC OF THE RECORD 1 SEC AFTER EACH INJECTION
(REF. 7)





All stretch-sensitive units have properties in common. All respond to stretch with a regular rhythmic discharge of impulses. The rate of firing is somewhat higher during and immediately after the imposition of stretch, but the discharge rate rapidly reaches a relatively steady level, which is maintained for hours if the muscle stretch is held constant. When the tension on the muscle is increased, the number of impulses per unit time increases, but not in a linear fashion. The firing rate is approximately directly proportional to the log of the applied muscle tension.

This relationship holds only within a limited range of applied tensions. At high tensions, the response falls short of the expected proportionality. Moreover, for different receptors, the slopes of the curves relating discharge rate to log tension are different. Discharge frequency is thus one way by which the receptor signals intensity of stimulus to the central nervous system.

If the monitored strand of fibers contains several axons supplying stretch receptors, another correlate of intensity becomes evident. As stimulus intensity is increased, the number of units responding increases. The thresholds of stretch receptors are distributed in accordance with a normal frequency curve, so that an increase in the intensity of the stimulus, i. e., muscle stretch, recruits additional units.

Receptor variation in threshold is due partly to true variance in sensitivity to stretch and partly to variance in location in the muscle, some receptors bearing more of the brunt of muscle stretch than others. Each of these receptors fires at a frequency which is determined by the extent to which the stimulus exceeds the threshold of the individual receptor. These two intensity-signaling variables - number of active units and frequency of unit discharge - account for the grading of reflex response to various degrees of stretch. It may be noted parenthetically that the relationships between stimulus intensity and receptor discharge outlined above are not unique to the stretch receptor. Similar relationships have been observed in a wide variety of receptors, including some involved in reflex regulation of visceral and vascular structures.



The muscle spindle and the Golgi tendon organ both respond to stretch. However, they respond differently to muscular events because of their different methods of incorporation into the muscular complex. The muscle spindles occur within the muscle, lying parallel to and between the muscle fibers. The tendon organs are found near the ends of the muscle, in the tendons. The operational effect of this difference is shown schematically in Fig. 3-9. In A, the muscle spindle is shown to be mechanically connected "in parallel" with the muscle fibers. Thus, when the muscle is stretched by an external force, the spindle responds with a pulse discharge. On the other hand, when the muscle is contracted by the application of a stimulus S to the muscle fibers, the spindle is silent. The tendon organ, represented in B, operates "in series" with the active fibers of the muscle. As such, it is equally susceptible to, but does not distinguish between, mechanical stretch applied by a passive pull on the muscle and that applied by active muscular contraction, both being actions that exert tension on the organ.

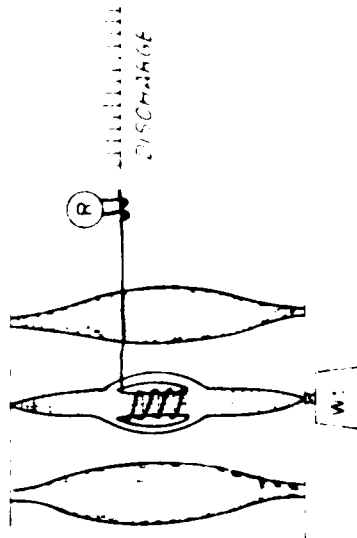
The muscle spindle is a fusiform body (i.e., spindle shaped, tapering at both ends). The spindle is located within the muscle itself, lying parallel to and between the muscle fibers. It is constituted of a bundle of from 3 to 10 muscle fibers (intrafusal fibers) enclosed in a fibrous capsule. The nerve supply of the spindle is double -- afferent and efferent.

An afferent fiber, upon entering the spindle and losing its myelin sheath and neurilemma, may end in one or other of two ways. (a) Some become flat and ribbon-like and wind in rings or spirals about the intrafusal fibers. These are called annulo-spiral or nuclear bag endings. (b) Others ramify upon the intrafusal fiber surface in a manner suggesting a spray of flowers. These are called flower-spray or myotube endings (Refs. 2 and 14).

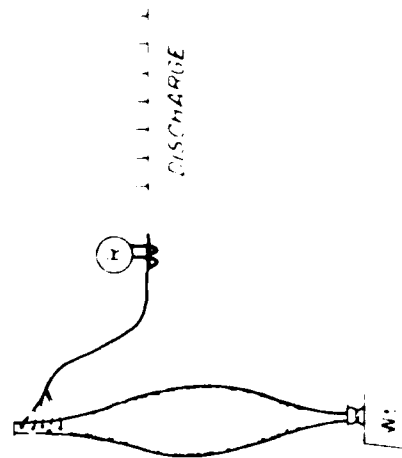
The efferent fibers, called fusimotor fibers or gamma efferents, terminate in end plates situated on the striated poles of the intrafusal fibres. According to Barker (Ref. 15), both poles of the intrafusal fiber receive this motor innervation. The function of these fusimotor fibers is apparently to furnish a variable threshold, or "bias", to the muscle spindles. It has been demonstrated that fusimotor activation causes contractions of the intrafusal fibers that are too feeble



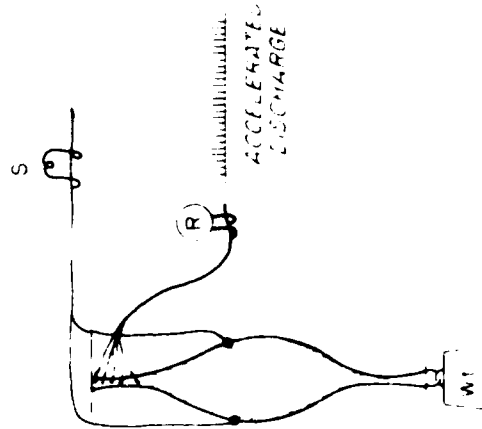
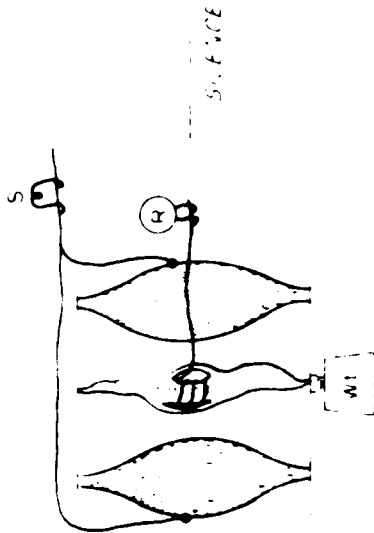
(A) SPINDLE



(B) TENDON ORGAN



MUSCLE STRETCHED



MUSCLE CONTRACTED

RELATION OF MUSCLE SPINDLES AND TENDON ORGANS TO MUSCLE FIBERS
(REF. 14)



to add significantly to the total muscle tension, but are sufficient to affect profoundly the afferent discharge of the spindle (Ref. 14). Whitteridge explored the behavior of the muscle spindles in the extraocular muscle of sheep and goats (Ref. 16). With the motor nerve cut and no stimulation to the gamma fibers, the spindles tested showed a linear relation between extension and frequency within the probable operating limits of the muscle. Sensitivities were found to occur in the range between 1 and 5 impulses/sec per millimeter of stretch. When the gamma fibers were stimulated a linear relation was still seen, but the sensitivities were plotted against frequency of the gamma stimulation (Ref. 16). The curves obtained were sensibly straight below 150 impulses/sec. Although gamma stimulation increases the sensitivity of the primary afferent endings of the muscle to spindle stretch, it appears to have little or no effect below their thresholds to stretch. However, there is evidence by Hunt (Ref. 17) that the threshold is reduced by stimulation of sympathetic nerve fibers, which differ from the gamma-efferent fibers and may be considered in the present context to be a third class of innervation.

Thus the muscle spindle provides a stretch receptor with variable sensitivity and threshold. The afferent discharge of the spindle depends not only upon stretch but also upon the number of impulses reaching the spindle via the fusimotor and sympathetic fibers. Each spindle receives up to five fusimotor fibers, providing a precise mechanism for grading the receptor sensitivity through a considerable range.

The Golgi tendon organ is characterized by a relatively very high threshold, usually requiring tensions of 100 to 200 grams or more for sustained firing (Ref. 14). It does not possess the efferent nerve fibers, and thus the variable gain and threshold characteristics of the muscle spindle.



3.2 Pressure and Touch

Pressure receptors may be considered informally as falling into three functional groups. The first group senses hydrodynamic pressures within the organism. For example, the baroreceptors of the carotid sinus, described in the previous section, monitor the blood pressure. As stated earlier, these are frequently classed as **stretch receptors** because of their operating mechanism and, as such, were treated in Section 3.1. The second group, of which the Lorenzian ampulla is an example, senses hydrostatic pressure outside the organism. There is evidence that this ampulla is employed by certain fish to sense their depth under the surface of the water. The third functional group senses what may be called contact-type pressure, e.g., that caused by pressing fairly hard against the skin. Although not necessarily restricted to it, the Pacinian corpuscle is an example of the third group.

Touch, or tactile sensation, is aroused by light contact (in contrast to the heavier contact associated with the pressure sensation). The tactile sensation is subserved in man by three types of receptor: Meissner's (Wagner's) corpuscles, Merkel's disks, and a basket-like arrangement of nerve fibers, surrounding the base of hair follicles. Meissner's corpuscles are situated in the papillae of the skin, just beneath the epidermis. They are unevenly distributed, being sparsely scattered over such a region as the volar aspect of the forearm but numerous in the skin of the hand, foot, nipple and lips, and in the mucous membrane of the tip of the tongue. They are well organized structures, consisting of irregularly coiled nerve endings with capsules of connective tissue. Merkel's disks consist of groups of three or more cup-shaped disks with a reticulated appearance. The nerve fiber upon approaching a group of such structures breaks up into branches, one going to each disk. Merkel's disks are found in the skin of the snouts of pigs and other mammals and in the finger-tips, lips and mouth of man. The basket-like arrangement, surrounding the base of a hair follicle consists of a number of short, vertical, nerve filaments which end in small bulbous expansions. They are stimulated by any slight movements of the hairs (Ref. 2).

Inequality of pressure with consequent deformation of the skin surface is the essential factor in the stimulation of touch receptors.



3.2.1 Pressure

Since those sensors described above as belonging to the first, or baroreceptor, group are represented in Section 3.1, they will not be discussed further here. Attention is directed first to the Pacinian corpuscle, and then to the less well understood Lorenzini corpuscle.

"The Pacinian corpuscles (Fig. 3-10) are oval bodies composed of concentric laminae, like the "skins" of a sectioned onion. The afferent fiber penetrates to the center of the corpuscle. Pressure with elongation of the organ and consequent stretching of the nerve ending is the adequate stimulus. These receptor organs are found in tendons, joints, periosteum, especially beneath tendinous insertions, in fasciae covering muscles and in subcutaneous tissues. They are also found in the mesentery. The structure of these receptors appears to be such that a mechanical stimulus such as stretching pressure, etc., induced by muscular action will be applied most effectively to the naked axis cylinder within its center" (Ref. 2).

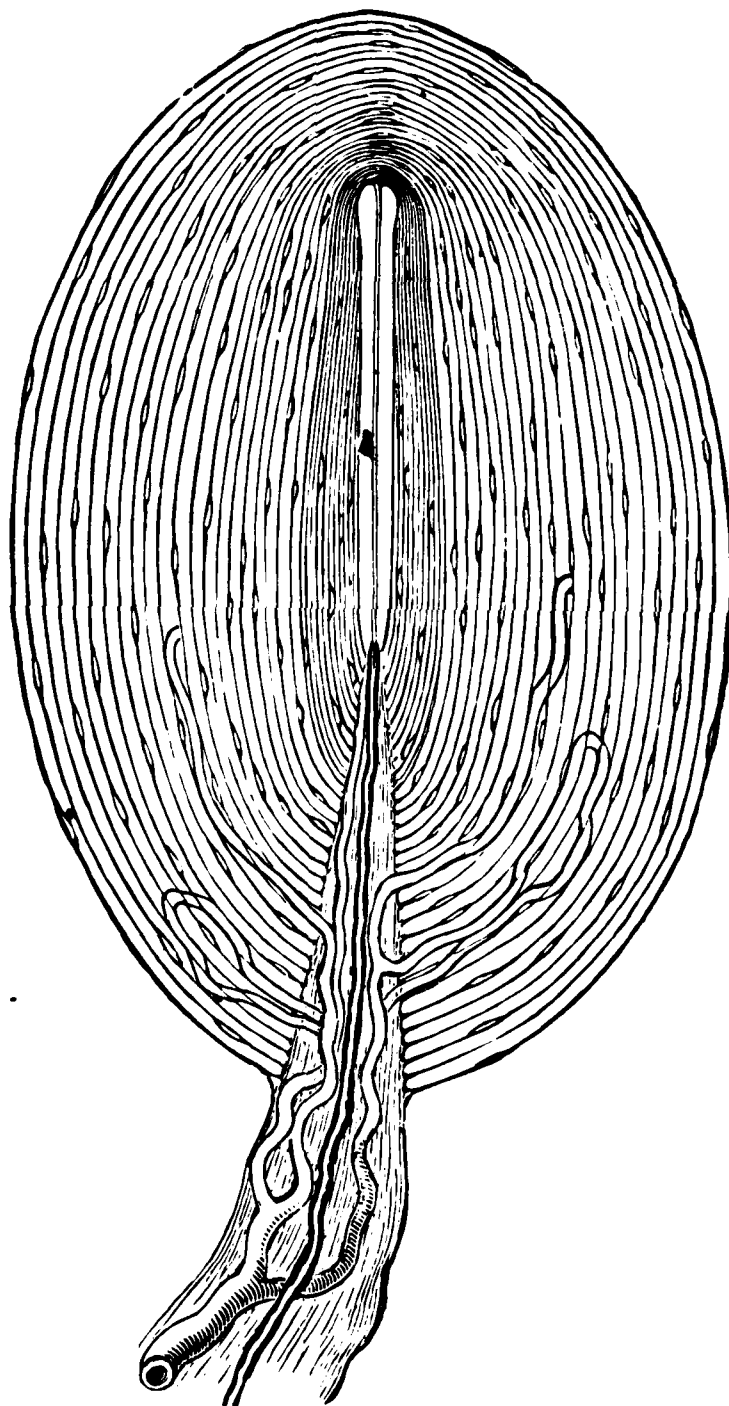
Morphological and histological data on the Pacinian corpuscle are quite abundant (e. g., Refs. 18 and 19). Several variations on the structural details have been found, involving principally branching and arborization of the axis-cylinder termination and the inner bulbs and lamellae.

The Pacinian corpuscle is described as rapidly adapting and sensitive to very light touch (Ref. 20). A deformation of 0.5 microns for 100 milliseconds will excite the corpuscle. Studies of the minimum value of extent and velocity of stimulus required for excitation showed a typical value to be 0.5 micron in 80 μ -sec which corresponds to a velocity of 0.006 m/sec. In experiments with short (<500 μ sec) repetitive mechanical pulse stimuli, it was found that the minimum interval between action potentials for any given stimulus strength agreed with values expected from direct measurements of refractory period (1 to 2 m/sec). In one case a stimulus frequency of 1000 cps was followed for many seconds. (Ref. 21).

Removal of the peripheral zones of the corpuscles appears not to impair their mechano-responsiveness (Ref. 22). It is proposed that the Pacinian corpuscle acts simply as a means of applying the mechanical stimulus to a section of the axon, and of providing appropriate mechanical damping. It is capable of protecting the axon from injury by the high pressures present at many of the corpuscle sites (Ref. 22).



PACINIAN CORPUSCLE
(REF. 1)





The initial application and the removal of the stimulus is much more effective than its steady maintenance in producing output pulses (Ref. 21). The corpuscle thus operates in part as a differentiator, or high-pass filter. It is believed not to be a mediator of constant pressure (Ref. 20). Although this property may be partly neural in origin, there is some evidence that it arises from the way in which the layers of the corpuscle transmit the mechanical stimulus to the core. Studies of the static and dynamic displacement of the corpuscle laminae (Ref. 23) indicate this kind of behavior.

The relative insensitivity to temperature of the Pacinian corpuscle is of interest from an engineering point of view. Eighteen corpuscles were subjected to temperatures ranging from 0° to 42°C (and in a few cases to 60°C) after it was determined that each was responsive to mechanical stimuli. In no case was there a response to temperature or temperature changes (Ref. 24).

The Lorenzinian ampulla is a large tubular organ that occurs in groups in the Elasmobranches. In the dogfish, for example, each ampulla consists of a bulb-like swelling about 0.7 mm in diameter, located deep within the head of the fish, and a jelly-filled tube several mm long that opens in a skin pore. The ampulla is innervated by 4-8 nerve fibers that end inside the swelling (Ref. 25).

Because earlier investigators found the ampullae to exhibit a high temperature coefficient (Refs. 26 and 27), they have been widely regarded as thermoreceptors. However Lowenstein (Ref. 25) believes that their deep location in the head and the relatively low thermal conductivity of the jelly argue strongly against their being thermoreceptors. He determined experimentally that when the internal pressure of the bulb was raised above the external pressure, the generator potential increment is sufficient to produce nerve impulses that vary in frequency with the pressure. When the internal pressure was abruptly increased, the nerve discharge had both a dynamic and a static non-adaptive phase. The frequency of both phases increases with pressure differential.

Lowenstein concludes (Ref. 25): "The Lorenzinian ampulla may thus possibly serve as a receptor of hydrodynamic pressure. In the dynamic phase of impulse discharge the sensitivity of the receptor at about 10 meters of water (the normal cruising range of the fish) is from 5 to 12 impulses/second per meter of water pressure"



3.2.2 Touch

Three types of touch sensor are listed in the introduction to Section 3.2: Merkel's disks, Meissner's corpuscles, and hair follicles. Each of these structures represents a way in which the cutaneous nerve fiber may terminate at the locale of the stimulus.

The nerves associated with general sensations, i.e., heat, cold, pain, pressure, touch, etc., are widely distributed throughout the body. They may end free among the tissue elements, or in special end organs where the terminal nerve filaments are enclosed in capsules.

"Free nerve-endings occur chiefly in the epidermis and in the epithelium covering certain mucous membranes; they are well seen also in the stratified squamous epithelium of the cornea, and are also found in the root-sheaths and papillae of the hairs, and around the bodies of the sudoriferous glands. When the nerve fiber approaches its termination, the medullary sheath suddenly disappears, leaving only the axis-cylinder surrounded by the neurolemma. After a time the fiber loses its neurolemma, and consists only of an axis-cylinder, which can be seen, in preparations stained with chloride of gold, to be made up of fine varicose fibrillae. Finally, the axis-cylinder breaks up into its constituent fibrillae which often present regular varicosities and anastomose with one another, and end in small knobs or disks between the epithelial cells.

"Under this heading may be classed the tactile disks described by Merkel as occurring in the epidermis of the pig's snout, where the fibrillae of the axis-cylinder end in cup-shaped disks in apposition with large epithelial cells." (Ref.1).

The special end organs include the Pacinian corpuscle, described in Section 3.2.1 and Meissner's corpuscles, of which the latter is a touch sensor. The special end organs have one feature in common; the terminal nerve fibrillae are enveloped by a capsule the function of which is to provide a proper match between the nerve ending and the stimulus so that the desired static and dynamic transduction characteristics will be obtained. In some cases the end organ protects the nerve ending from irreversible damage from overload.



"The tactile corpuscles of Wagner and Meissner are oval-shaped bodies. Each is enveloped by a connective-tissue capsule, and imperfect membranous septa derived from this penetrate the interior. The axis-cylinder passes through the capsule, and after making several spiral turns around the body of the corpuscle ends in small globular or pyriform enlargements. These tactile corpuscles occur in the papillae of the corium of the hand and foot, the front of the forearm, the skin of the lips, the mucous membrane of the tip of the tongue, the palpebral conjunctiva, and the skin of the mammary papilla." (Ref. 1).

Microphotographs of the Meissner's corpuscle and related tissue in the papillary ridges of the finger are included in Reference 28. In particular, a rubber model of the papillary ridge is shown. This model demonstrates the way in which the ridge construction acts as a magnifying lever mechanism for transmission of the touch stimuli to the underlying receptors. The reference states that Meissner's corpuscle is fully stimulated by pressure coinciding with its axis.

The touch sensor remaining to be discussed is the hair follicle. All hair follicles are innervated and, therefore, are tactile in function (Ref. 29). Of particular interest, however, is the tylotrich follicle because of its specialized construction and high sensitivity.

Some mammals, such as the cat and rabbit, possess guard hairs. In general these may be differentiated from pelage hairs which are shorter, finer and more dense than the longer, coarser and more sparsely located guard hairs. These hairs may be divided into three general types: vibrissae, common guard hairs and tylotrichs. Vibrissae are most commonly found around the nose, chin and eyes of fur bearing mammals. These act as aperture sensing receptors, i.e., they judge the width of a passage through which the animal may wish to pass. The common guard hairs are found over the surface of the animal and serve a two fold function: protection of the soft pelage from environmental damage and mediation of the tactile sense.

The tylotrichs are longer than common guard hairs; they are very sparse, are straighter than the common guard hair, and they have longer, more gradually tapered tips. Their follicles are profusely innervated and contain erectile tissue,



i. e., tissue containing capillaries that dilate after stimulation. A dense network of nerve fibers resides between the erectile tissue and the external root sheath. Thus, erection of the annulus and possible contraction of the band of smooth muscle-like cells provides pressure on the nerve ending and consequent mechanical amplification of the stimulus. The tylotrich follicle, then, has a high and variable sensitivity (Ref. 29).

3.3 Acceleration Receptors

In the present context, the term acceleration receptor will be taken to include sensors of linear acceleration, gravity gradient, and angular rotation and acceleration. The biological devices to be discussed in this section are

statocyst
semicircular canal and utricle
lateral line organs
halteres.

A sense of spacial orientation is vital to living organisms. Even plants require that their roots grow down into the earth and their shoots grow up into the air and sunlight. Thus, virtually all living forms possess some form of spacial sensor, rudimentary though it may be, to control their posture through either motor control or directed growth. The requirement for this orientation arises basically from the efficiency of the nourishment process, and to a lesser degree, perhaps, from consideration of defense of the organism.

Free-moving animals require something more than simple attitude control. The complexity of the need, in any given case, is determined by the characteristics of the body and of the locomotive process. Locomotion is basically concerned with feeding and defense. Examples of the more sophisticated postural and locomotive systems are found in the higher quadrupeds, man, and in creatures of flight. Dynamic stability and high maneuverability are essential elements in the competition for survival, serving both the feeding and escape functions. Dynamic stability requires the sensing and appreciation of velocity and acceleration, in addition to simple position and attitude. The evolutionary process has developed a variety of organs suitable for these measurements.



3.3.1 Statocyst

The statocyst organ is most commonly found in molluscs and crustacea. It may be thought of as a fluid-filled sphere whose lining is characterized by sensory epithelium localized either on the top or bottom (Ref. 30). The epithelial cells frequently possess hair-like projections that protrude into the chamber. A solid or semisolid body, the statolith, rests on or hangs from this sensory epithelium. The statolith, which is usually a calcium salt or sand grain, provides the pendulous mass for the sensor; it triggers nerve discharges according to its position relative to the sensory epithelium. Any change in position of the statolith, whether caused by linear or angular displacement, causes a differential pressure gradient on individual epithelial cells. This causes impulses to be sent to the coordination centers of the central nervous system, where an appropriate motor response is initiated.

An interesting experiment performed by Kreidl is described in Reference 1. A crustacean that had just undergone ecdysis was obtained. Characteristically, when the shell was shed, some of the inner surface of the lateral portion of the statocyst was destroyed. The animal was placed in an isolated area where only iron filings were present— not the usual sand or other particles. The result was that the iron filings were picked up and incorporated into the statocyst to form the new statolith. It was then found that the orientation of the animal could be directly influenced by the application of externally applied magnetic fields.

An extensive study of the statocyst in the lobster Homarus americanus has been undertaken by Cohen (Ref. 31). By studying the response of the organ, he identified four types of receptors according to the nature of their response:

- Position receptor, Type 1
- Position receptor, Type 2
- Acceleration receptor
- Vibration receptor.

Position receptor, Type 1—These receptors show a specific non-adapting discharge frequency for different maintained positions about the transverse axis. They are sensitive primarily to absolute position. The receptor threshold near



the region of maximum response appears to be less than 6° of displacement.

Position receptor, Type 2—These receptors are capable of signalling absolute position in a manner similar to that described above. In addition to this static behavior, there is a differentiated dynamic response which seems capable of signalling also the direction from which a given position is approached.

Acceleration receptor—Specific receptors respond only to motion about the horizontal and vertical axes, not to position. Although the article was not as precise as it might have been in distinguishing rotational acceleration from velocity, it appears that acceleration is the sensed parameter since the modulation of impulse frequency occurs at the start and termination of the angular motion. The threshold for this effect is quoted as occupying the range between 5 and $15^\circ/\text{sec}^2$. The impulses return to resting level within 0.5 sec after constant angular velocity is reached.

Vibration receptor—Certain receptors were found that respond only to vibration conducted through the substrate—not through the surrounding air or water.

Cohen determined that removal of the statolith abolishes all position responses but does not affect the acceleration or vibration receptors. Removal of both the statolith and the fluid from the cyst abolished all responses except that to vibration (Ref. 31). One is tempted to infer from this that the pendulus mass (statolith) operates the position sensors, that the fluid motion itself supplies sensitivity to angular acceleration, and that the epithelial cells and their hairs respond to vibratory stimuli.

The properties of the Type 2 position receptor suggest an interesting area for further research. The ability to determine the direction from which a position is approached indicates some sort of memory and logic scheme in the central nervous system. To one interested in the general area of neural behavior and data processing, with applications for example to pattern recognition, every possible occurrence of this mechanism in nature affords a potential opportunity for fundamental study.



3.3.2 Utriculus and Semicircular Canals

Man and other vertebrates derive position and motion data from the inner ear, or labyrinth. The labyrinth consists of two parts: the osseus labyrinth, a series of cavities in the temporal bone, and the membranous labyrinth, a structure of communicating membranous sacs and ducts contained within the bony cavities. The membranous labyrinth is illustrated in Figure 3-11. It roughly corresponds in form and shape to the osseus labyrinth and is contained in it; it is necessarily somewhat smaller.

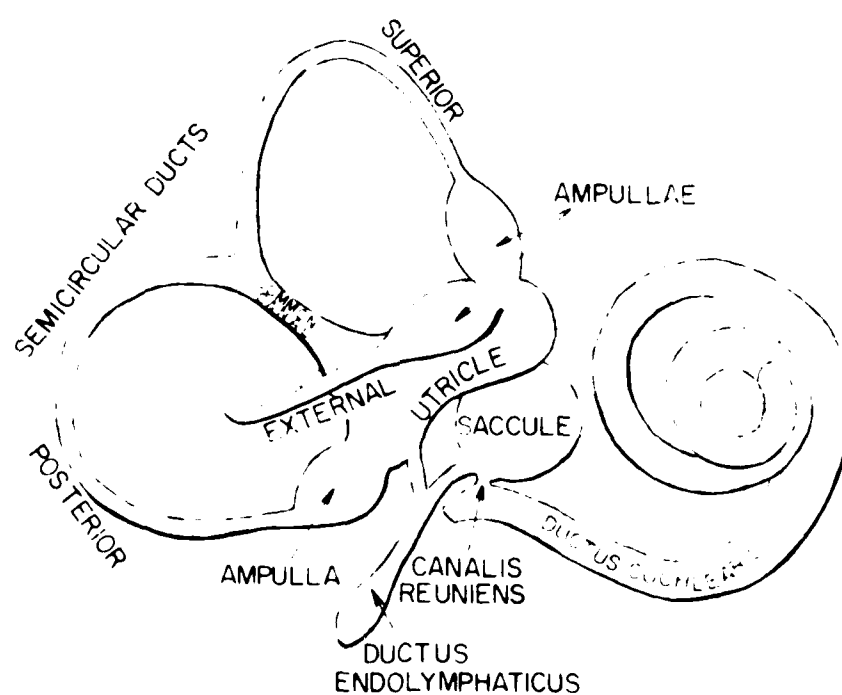
The osseus labyrinth consists of three parts: the vestibule, semicircular canals, and cochlea. The components of the membranous labyrinth that correspond to these divisions are the utricle (utriculus) and saccule in the vestibule, and the semicircular and cochlear ducts in the respective canals. Details of construction of the labyrinths may be found in Reference 1.

The three semicircular canals lie in planes that are very nearly mutually perpendicular. Each canal, in conjunction with the utricle, forms a closed-loop path for the fluid contained therein. An ampulla in each duct (Fig. 3-11) senses flow of the internal fluid relative to the duct. When the organ is rotated in the plane of one of the ducts, the inertia of the fluid column tends initially to hold it stationary, so that it possesses a relative negative rotation with respect to the duct wall. This relative rotation is sensed by the ampulla. If the rotation of the organ continues at a constant angular rate, the fluid inertia is overcome by frictional and capillary forces so that eventually the fluid and duct rotate together. At this time, there is no relative motion for the ampulla to detect. If rotation of the organ should be suddenly arrested, the fluid would continue to rotate for a time from its own inertia; the ampulla would then register a relative rotation in the direction opposite from that registered during acceleration. Thus, the semicircular canals and their ampullae comprise rotational acceleration sensors in three mutually perpendicular planes.

Detection of the relative fluid motion is accomplished by the cristae, the receptor organs of the semicircular canals. They are situated in the ampullae and consist of a structure of hairs and hair cells arranged to be sensitive to motion of the canal fluid. The detection threshold for angular acceleration is between 1 and $2^{\circ}/\text{sec}^2$ (Ref. 2).



THE MEMBRANOUS LABRINTH (REF. I)





The possibility that rectilinear acceleration is also detected by the semicircular canals has been proposed by McNally and Tate (noted in Ref. 2). The mechanism for this operational mode is not obvious since the ducts are of capillary size and are well supported. McNally and Tate apparently ascribe this rectilinear sensitivity only to the vertical canals.

The utricle and saccule are fundamentally similar in construction to the statocyst described in Section 3.2.1 and as such should be inherently capable of sensing gravitational pull and linear acceleration. Evidence concerning the function of these components is conflicting. However, sensitivity to both tilt and linear acceleration are credited to the utricle, while the present consensus is that the saccule is an organ associated with the cochlea and designed for the reception of slow vibrational stimuli (Ref. 14). It may be an organ for the registration of vibrations in the head transmitted through bone—as distinct from vibrations in air. The appreciation of the sound of one's own voice may possibly depend therefore upon saccular as well as upon cochlear function (Ref. 2).

3.3.3 Lateral Line Organs

The lateral line was once thought to contribute to the maintenance of equilibrium (Ref. 32). However, no information has come to light during the present investigation regarding any equilibration mechanism in the lateral line. The consensus seems to be that this organ is associated primarily with the reception of vibration; its description is therefore deferred to Section 3.4 which deals with that subject.



3.3.4 Halteres of Diptera

An interesting gyroscopic device is used for aerodynamic stability by the true flies of the order Diptera. It appears to give the insect sensory indications of rotation in the yawing plane. This capability is manifest in a pair of organs known as halteres. The information given here is taken from Pringle's work (Ref. 33).

The halteres are essentially dumbbell shaped organs that oscillate rapidly in vertical planes during flight. They are swept back from the transverse axis of the insect by an angle of approximately 30° . Their gross morphology varies from family to family according to the extent of evolutionary development; it is generally accepted that the halteres are modified hind wings. In the more primitive forms, such as Tipuls of the family tipulidae, the halteres are very long and exposed; in the more specialized cases they are very short (0.7 mm in Calliphora), and in the calyptrate Diptera are covered by a flap of the wing so that they oscillate in an enclosed space and are presumably little affected by the airstream.

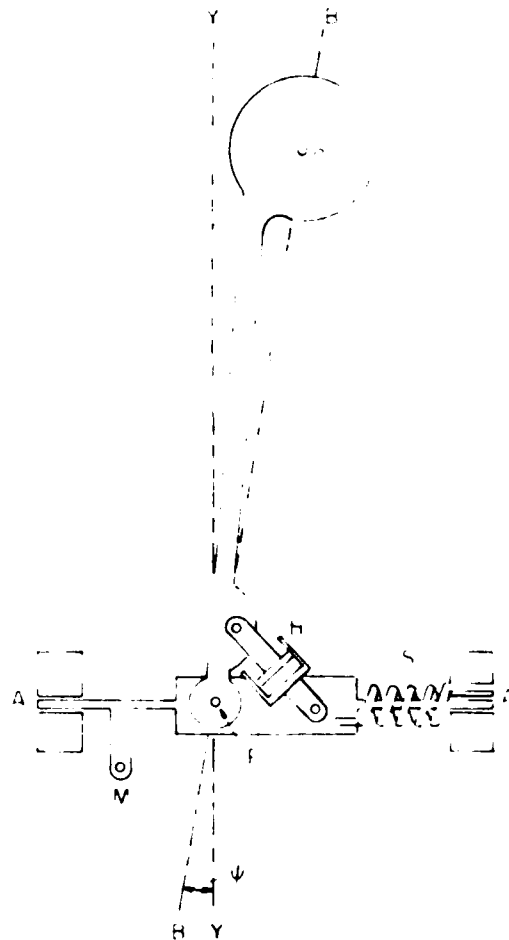
The construction of the haltere is exceedingly complicated both with respect to its articulation and to its sensilla. Since these are covered in detail in Reference 33, the present description will be restricted to that necessary for functional understanding.

A mechanical picture of the haltere operation may be gained by the analog diagram by Pringle (Ref. 33) and shown in Figure 3-12. The axis of oscillation is shown by the bearings AA. This oscillatory motion is sustained by a single muscle, connected at M and operating in a direction perpendicular to the paper, and by the restoring torsional spring S. The mass of the system and the elasticity of the spring determine the natural period of alternation of the haltere, to which successive contractions of the muscle must be adjusted in order that the amplitude of the oscillations build up. Sensors located at the base of the stalk serve to measure the amplitude of oscillation and regulate it through reflex inhibition of muscle contraction.

The secondary degree of freedom of the haltere is represented by the pin at P; damping of the motion about this axis, accomplished by the fluid contents of



DIAGRAM TO ILLUSTRATE THE MECHANICS OF THE
HALTERE ARTICULATIONS
(REF. 33)





the basal structure, is represented by the dashpot H. "When the haltere oscillates about the axis AA the mass of the moving portion, considered as located at the center of gravity C, moves round the circumference of a circle. Since the line BB through P and C makes an angle of less than 90° with the axis AA, there will be a torque tending to rotate the line BB into the position YY at right angles to AA. This torque is discontinuous, but it is always in the direction tending to rotate the line BB into the position YY; it fluctuates at a frequency twice that of the oscillation of the haltere. Provided the articulation at P is sufficiently free and the damping produced by the fluid contents of the haltere is sufficient to average out this fluctuating torque, the slight secondary freedom of movement will therefore ensure that the centre of gravity of the moving portion of the haltere always moves in a plane through P at right angles to the axis AA, i.e. a vertical plane if AA is horizontal. Curvatures and irregularities in the stalk are therefore unimportant, and when the haltere is oscillating with the insect at rest, the forces acting on the basal region just distal to the secondary articulation will always be in the vertical plane.

"Now consider the effect on the dynamics of the haltere of angular rotations of the entire fly. The angular momentum of the oscillating mass, being rotated out of its own plane of oscillation, will produce gyroscopic torques in the same way as if the mass were continuously rotating, except that these torques will not be constant in magnitude or direction. In order to obtain a satisfactory theoretical treatment of the problem of an oscillating gyroscope it is necessary to make three approximations, each of which will be examined later: (a) that the oscillating system distal to the articulations already described is rigid, (b) that the mass of the system is concentrated at a point, and (c) that the oscillation is a simple harmonic motion. With this simplification the properties of the system may be analyzed by the methods of dynamics. The analysis shows that, apart from the primary torques inherent in the oscillatory motion and present when the fly is stationary, angular rotations of the whole fly set up a torque at the base of the haltere acting about an axis at right angles to the moving stalk in the plane of oscillation; and, further, that whatever the plane of rotation of the fly this is the only new torque set up. The magnitude and periodicity of the torques at the base of the haltere differ for different planes of rotation." (Ref. 33).



As a result of his anatomical studies, experiments, and dynamic analysis, Pringle advances the following hypothesis to explain the function of the halteres:

- (1) "The halteres of Diptera are organs of special sense giving an indication to the fly of rotations in the yawing plane.
- (2) "The discharge of impulses in the haltere nerve is interpreted as indicating a turn to the left or to the right according as the impulses from the basal plate sensilla precede or follow those from the large chordotonal organ.
- (3) "The impulses from the dorsal and ventral scapal plate sensilla serve to measure the amplitude of the haltere oscillation and, by reflexly inhibiting the contractions of the haltere muscle, preserve a constant amplitude of oscillation.
- (4) "The rate of rotation in yaw is indicated by the relative timing of the impulses from the basal plate sensilla those from the large chordotonal organ, and also by their intensity (i. e. number of sensilla excited). If there is no separation in time, a simultaneous increase in intensity of both will have no effect. Slight maladjustment of the organ so that both groups are excited in the absence of rotation will therefore not be interpreted as a turn.
- (5) "For a given direction of rotation the two halteres will be oppositely excited; the indications must therefore be interpreted in opposite sense.
- (6) "Pitching and rolling rotations are not interpreted as turns owing to differences in the timing of the impulses. Since the haltere system is fundamentally incapable of distinguishing between pitching and rolling rotations, it is difficult to see how it can give useful indications in these planes." (Ref. 33).



In general aerodynamic terms, the fly probably shows what is known as "spiral instability". Any slight inequality in the force produced by wings of opposite sides will generate a rolling and yawing moment which, if not immediately corrected by some reflex mechanism, will lead to a spin (Ref. 33). The evidence suggests that the halteres mechanism provides the means for such a reflex correction.

The gyroscopic system represented by the halteres mechanism is striking in its elegance. However, its most interesting feature from the point of view of this study may be the form in which its data are presented to the insect. The fact that yaw rate is represented in part by the relative timing of pulses suggests that the insect has a capability to process information with pulse-position coding. This implication, together with the very small size of the processing system, makes the associated neural networks appear fully as interesting as the mechanical features of the gyroscope itself.

3.4 Sound and Vibration

The distinctions between sound and vibration are not universally drawn, nor will an attempt be made here to delineate them. If a rule-of-thumb is desired, it is convenient to separate sound from other vibratory phenomena on the basis of the transmitting medium, i.e., to say that vibrations transmitted through the air are sound. Such a distinction is adequate for the present purposes.

3.4.1 The Ear

Descriptions of the anatomy and physiology of the ear of the human and many other species abound in the literature (e.g., Refs. 1, 2, 14, 34, through 38). Therefore only a general description and some points of interest are presented here. The human ear will serve as the basis for these remarks although its fundamentals are shared with the auditory organs of many species.

The ear may be considered functionally to consist of two main elements: a transducer and an impedance-matching device. The transducer, which converts the sound waves into nerve impulses is called the cochlea. The remainder of the auditory mechanism, consists of the external ear (auricula),



eardrum (tympanum), auditory canals, and a delicate sequence of bony components. The function of these latter components is to match the acoustic impedance of the cochlea to the ambient air, thereby transmitting the pressure variations from the air to the cochlear fluid with little energy loss.

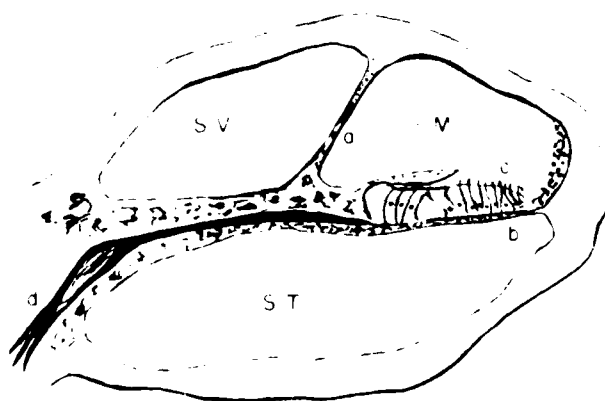
The cochlea is a spiral canal which is described by Best and Taylor as follows:

"The canal makes $2 \frac{3}{4}$ turns round a central pillar of bone called the modiolus. The smallest turns of the cochlea are at its apex. A ledge of bone winding around the modiolus like the thread of a screw-nail divides the spiral canal incompletely into two parts. The partition is completed by a membranous structure—the basilar membrane—which extends from the tip of the lamina spiralis ossea to the center wall of the canal. A second membrane—Reissner's membrane—stretches from the upper surface of the bony lamina to a point a short distance above the outer attachment of the basilar membrane. The original osseous canal is divided in this way into three spiral compartments or galleries (Fig. 3-13). The gallery below the basilar membrane is called the scala tympani, the one above Reissner's membrane, the scala vestibuli, while the one enclosed between the two membranes is known as the scala media, the cochlear duct or the membranous cochlea. The membranous cochlea is filled with a fluid known as endolymph, while the osseous canals, i.e., the scala vestibuli and scala tympani, are filled with perilymph. The scala vestibuli and scala tympani communicate with one another at the cupula of the cochlea through a small aperture called the helicotrema. The scala media ends here as a closed sac, but its basal extremity communicates with the non-auditory labyrinth through the canalis reuniens or duct of Hensen."

"The auditory receptors. The sensory cells together with sustentacular elements constitute a structure known as the spiral organ of Corti. This lies within the scala media, occupying the inner half or so of the basilar membrane. It presents towards its inner part two rows of elongated epithelial elements of peculiar shape—the inner and outer rods (or pillars) of Corti—which stand with their expanded bases planted upon the basilar membrane. The bases of the rods



CROSS SECTION OF COCHLEAR CANAL (REF. 2)



- SV - STALA VESTIBULI
- VM - STALA MEDIA COCHLEAR DUCT
- ST - STALA TYMPANI
- a - REISSNER'S MEMBRANE
- b - BELLER MEMBRANE
- c - ORGAN OF CORTI
- d - AUDITORY NERVE



of the two rows are separated by an interval, but above the membrane their bodies incline towards one another and, meeting at an angle, their extremities fit into one another in a manner much as one clasps a fist with the other hand. They form by this arrangement, a triangular tunnel whose floor is covered by two nucleated scraps of protoplasm which actually are the undifferentiated parts of the cells from which the corresponding rods have developed. The arches of Corti increase progressively in height from the basal to the apical turns of the cochlea.

"The hair cells are the essential sensory elements. They are seen as four rows of short columnar cells, one row lying on the inner side and three on the outer side of the corresponding rod. The inner hair cells number about 3500; their free ends are in line with the surfaces of the rods, but their bases extend only half way to the basilar membrane. The free surface of each hair cell of both inner and outer rows is surmounted by some twenty hair-like processes. One or two rows of cells support the bases and inner aspects of the inner hair cells. The outer hair cells are longer than the inner and more numerous (12,000); they are separated from the basilar membrane by cells arranged in several rows—the cells of Deiters—which send slender processes between the rows of the sensitive cells. That part of the organ of Corti lying on the outer side of Deiters' cells and the outermost row of hair cells is composed of several layers of columnar cells—the supporting cells of Hensen.

"The tectorial membrane is a delicate, almost homogenous structure, somewhat paddle-shaped in transverse sections of the scala media; it overlies the inner half or more of the spiral organ. Arising from a point near the base of the lamina spiralis ossea it ends laterally in an irregularly fringed or scalloped free margin. The hairs of the sensory cells of all four rows are embedded in its under surface." (Ref. 2).

There are various theories of the manner in which sound waves are converted to nerve impulses, and subsequently to the sensations of sound. In particular, the ability to discriminate both intensity and pitch is not firmly understood. The theories may be divided into:



1. Place theory

- a. resonance
- b. traveling wave

2. Frequency theory

The place theories involve the manner in which the basilar membrane responds to different sound frequencies; neither specifies the mechanism of end-organ excitation. The frequency theory does not require the cochlea to respond differentially along its extent for frequency reception. Rather, the frequency of the sound wave is presumed to be reproduced in the frequency of nerve discharge, the cerebral cortex providing the interpretation function. There are many difficulties to the frequency theory and the evidence for the place theory seems overwhelming (Ref. 14).

The cochlea represents the most sensitive of the general class of mechanoreceptors (Ref. 34). The minimum audible energy for the human ear is about 10^{-12} erg, compared to the minimum visible energy of about 10^{-10} erg (Ref. 35). Its sensitivity, plus the sophistication of its sensing and data processing mechanism recommend it for considerably more study in any continuation of the present survey.

3.4.2 Weberian Ossicles and Swimbladder

The minnow is endowed with a mechanism which is believed to have among its functions the sensing of vibrations. The swimbladder is an elongated sausage-shaped organ of two parts—the anterior and posterior sacs. The organ is located in the abdominal cavity dorsal to the alimentary canal and ventral to the vertebral column. At the anterior end of the organ there is a gland, or red body, which maintains proper inflation of the sac by the secretion or reabsorption of gas.

Although the swimbladder is considered to be associated with the measurement of pressure, there are features of it that suggest its capability as a vibration sensor. Evans (Ref. 39) describes a connection between the swimbladder and the ear as follows:

"The most characteristic feature of this group (family Cyprinidae) is the modification of the anterior four vertebrae, often co-ossified and bearing a chain



of small bones, the Weberian ossicles, which connect the air bladder with the ear. The most posterior of these ossicles is attached to the anterior end of the anterior sac. The physiological significance of these so-called ossicles is still an unsettled problem." (Ref. 39).

Prior to 1915, biologists did not share a common view on the subject and work was being done on both pressure and vibration sensing. However, new experimental investigations in 1915 brought new facts to light. Kuiper destroyed the Weberian ossicles and noted the reaction of fish exposed to pressure changes, light alternations, vibrations, and static disturbances (discussed in Ref. 39). The fish in which the ossicles were destroyed reacted to all stimuli except vibration. He then experimented with fish in which the eighth cranial nerve was cut and found that the fish would respond to light changes but not to vibration or static disturbances.

Subsequent thinking seems to reflect Kuiper's results. Evans (Ref. 39) states that, "The arrangement of ossicles seems peculiarly designed for the purpose of recording rapidly-recurring movements of small amplitude." He also states a second point in support of the vibration capability: "The wall of the air bladder is capable of vibrating synchronously with rapidly-recurring sound waves."

The question remains unresolved, but the evidence for a vibration-sensing function in the swimbladder is substantial.

3.4.3 Lateral Line Systems

Lateral line systems are found in the cyclostomes, fishes, aquatic larval stages of amphibians, and in various degrees of development in some adult amphibians. In most fish, the system comprises a series of mucous-filled canals enclosing sense organs and opening to the surface at intervals through small pores. These canals generally branch considerably in the cephalic region. There is considerable variation within the group. Some fish have superficial organs in addition to those in the canal; others have organs located in shallow surface grooves rather than in the deeper canals. Amphibians exhibit this superficial type of sense organ arranged in a series of lines.



The general structure of the sense organs, or neuromasts, associated with the lateral line systems is pear-shaped in cross section. The sense cells are centrally located and are tear-shaped, tapering toward the free surface of the organ. Each cell bears a hair-like process at its tip. There is a large ovoid nucleus in each sense cell, and mitochondria appear to accumulate at the distal end. The sense organs are innervated and supplied with capillaries through apertures in the scales.

The canals of the lateral line are filled with jelly-like mucous and are always in contact with the surrounding water. The fluid in the canal of the swimming fish, evoked by the flow of the external medium through the lateral line pores, produces a ciliary bending of the sensory cells. This bending, combined with the deformation of these cells, stimulates the production of nerve impulses.

The lateral line system is found to be sensitive to vibratory phenomena of frequencies up to 344 cps (Ref. 36). It is theorized that the lateral line organs are receptors in a short-range auditory system (Ref. 35).

3.4.4 Lyriform Organ

It has been shown that at least two species of spider have sensitive vibration receptors located on each of their 8 legs (Ref. 40). The spider uses this sense to detect, and to some extent identify, objects that come in contact with its web. His behavioral response to a vibration stimulus depends upon the frequency and duration of the stimulus and may consist of either attack or escape.

The receptor is the lyriform organ, of which a detailed description may be found in Reference 41. It is extremely sensitive to vibration. Walcott and Van der Kloot (Ref. 40) determined that at the frequency of its peak sensitivity the receptor responds to a mechanical displacement as little as 2.5×10^{-6} mm. It was demonstrated that the spider is capable of responding to web vibrations at frequencies from 400 to 3,000 cps, and that it can distinguish vibrations at 400-700 cps from those at frequencies above 1,000 cps. The receptor is also



sensitive to airborne sound. Under ideal conditions it will respond to a 4,000 cps signal at a sound pressure of 23 db -- the threshold of the human ear at this frequency is about 3 db (Ref. 40).

The vibration receptor responds to increasing stimulus intensity with an increase in the rate of nerve impulses. The organ adapts to a continued stimulus in 20-30 sec, and requires about 10 minutes to completely recover from the adaptation. The speed of response to a newly-presented stimulus is in the range of 0-15 ms depending on frequency and intensity of the stimulus.

3.5 Conclusion

This section on mechanoreceptors represents a beginning of a survey of the field, employing the "cataloging" approach discussed in Section 2. It is apparent that additional effort could be profitably expended in filling out the areas covered, as well as extending the study to new areas.

The ear represents an example of a sensor about which a great deal of literature has been written and which offers some interesting possibilities in at least two directions. First of all, the ear is a rather elegant acoustic device. Second, the ear and its associated neural system is a living information-processing system, the operation of which is far from being well understood. The variations in form and complexity of aural receptors among the phyla offer a broad source of information from which both the acoustic and neural operations may ultimately be deduced.

Another point about mechanoreceptors involves a potential application. The Navy has for some time been interested in the way that fish are able to swim at widely varying speeds without encountering the problems of cavitation that plague boat designers. The optimum shape of a body moving through the water is highly dependent on the speed of the body. It appears that fish are able to adjust their shape and manner of locomotion to a degree sufficient to meet the cavitation problem. This suggests a reflex system in which sensors are able to detect incipient cavitation conditions and effect a motor response to correct the situation. It seems quite natural to expect that a continuation of the present study into areas relevant to this question could accelerate a solution to the problem. This case is one example typical of the potential yield of applicable material within this category which we have not as yet fully explored.

REFERENCES SECTION 3

1. Gray, H., Anatomy of the Human Body, 25th Edition, Lea and Feliger, Philadelphia, (1948).
2. Best, C.H. and Taylor, N.B., The Physiological Basis of Medical Practice, 5th Edition, Williams and Wilkins Co, Baltimore, (1950).
3. Kezdi, P., Control by the Superior Cervical Ganglion of the State of Contraction and Pulsatile expansion of the Carotid Sinus Arterial Wall, Circulation Research, Vol. 2, p. 367, (July, 1954).
4. Bronk, D.W. and Stella, G., The Response to Steady Pressures of Single End Organs in the Isolated Carotid Sinus, American Journal of Physiology, Vol. 110, p 708, (1935).
5. Ead, H.W., Green, J.H. and Neil, E., A Comparison of the Effects of Pulsatile and Non-Pulsatile Blood Flow Through the Carotid Sinus on the Reflexogenic Activity of the Sinus Baroreceptors in the Cat, Journal of Physiology, London, Vol. 118, pp 509-519, (1952).
6. Warner, H.R., The Frequency-Dependent Nature of Blood Pressure Regulation by the Carotid Sinus Studied With an Electric Analog, Circulation Research, Vol. 6, p. 35, (January, 1958).
7. Paintal, A.S., A Study of Right and Left Atrial Receptors, Journal of Physiology, London, Vol. 120, pp 596-610, (1953).
8. Coleridge, J.C.G., Hemingway, A., Holmes, R.L., Linden, R.J., Atrial Receptors in the Dog, Proc. Physiological Soc. 20-21, p 68P, (April, 1956).
9. Henry, J., Gauer, O., Reeves, J., Evidence of the Atrial Location of Receptors Influencing Urine Flow, Circulation Research, Vol. 4, p. 85, (1956).
10. Nonidey, J.F., Identification of the Receptor Areas in the Venae Cavae and Pulmonary Veins Which Initiate Reflex Cardiac Acceleration, Bainbridge's Reflex, American Journal of Anatomy, Vol. 61, p. 203, (1937).



11. Henry, J.P. and Pierce, J.S., The Possible Role of Cardiac Atrial Stretch Receptors in the Induction of Changes in Urine Flow, *Journal of Physiology*, London, Vol. 131, p. 572, (1956).
12. Gauer, O., Henry, J., Speiker, H., and Wendt, W., The Effects of Negative Pressure Breathing on Urine Flow, *Journal of Clinical Investigation*, Vol. 33, p. 287, (1954).
13. Page, Hickam, Speiker, McIntosh, and Pryor, Reflex Venomotor Activity in Normal Persons and in Patients with Postural Hypertension, *Circulation*, Vol. 11, p. 262, (1955).
14. Ruch, T.C., and Fulton, J.F., Medical Physiology and Biophysics, 18th Edition, W.B. Saunders Company, Philadelphia, (1960).
15. Barker, D., The Innervation of the Muscle Spindles, *Quarterly Journal of Microscopic Science*, Vol. 89, pp 143-186, (1948).
16. Whitteridge, D., The Effect of Stimulation of Introfusal Muscle Fibers on Sensitivity to Stretch of Extraocular Muscle Spindles, *Quarterly Journal of Experimental Physiology*, Vol. 44, p. 385, (1959).
17. Hunt, C.C., The Effect of Sympathetic Stimulation on Mammalian Muscle Spindles, *Journal of Physiology*, London, Vol. 151, pp 332-341, (1960).
18. Takashi, M., Sokai, I., Usiyina, H., On the Terminal Neural Apparatus Detectable in the Retroperitoneum of Man - A Complex Pattern of Pacinian Corpuscle, *Anatomical Record*, Vol. 122, p. 17, (1955).
19. Pease, D.C., Quilliam, T.A., Electron Microscopy of the Pacinian Corpuscle, *Journal of Biophysical and Biochemical Cytology*, Vol. 3, No. 3, p. 331, (1957).
20. Gray, J.A.E. and Matthews, P.B.G., Response of Pacinian Corpuscles in the Cat's Toe, *Journal of Physiology*, London, Vol. 113, p. 475, (1951).
21. Gray, J.A.B and Malcolm, J.L., The Initiation of Nerve Impulses by Mesenteric Pacinian Corpuscles, *Proc. Royal Society, London*, Vol. 137, p. E96, (1950).



22. Loewenstein, W.R., Rathkamp, R., The Sites for Mechano-electric Conversion in a Pacinian Corpuscle, *Journal of General Physiology*, Vol. 4, No. 6, p. 1245, (1958).
23. Hubbard, S.J., The Mechanical Behavior of Pacinian Corpuscles, *Proc. Physiological Society London*, Vol. 26-27, p. 40P, (April, 1957).
24. Loewenstein, W.R., On the 'Specificity' of a Sensory Receptor, *Journal of Neurophysiology*, Vol. 24, p. 150, (1961).
25. Loewenstein, W.R., Mechanisms of Nerve Pulse Imitation in a Pressure Receptor (Lorenzinian Ampulla), *Nature*, Vol. 188, No. 4755, p. 1034, (1960).
26. Sand, A., The Function of the Ampullae of Lorenzini With Some Observations on the Effect of Temperature on Sensory Rhythms, *Proceedings of the Royal Society*, Vol. 125, p. 524, (1938).
27. Hensel, H., Die Wirkung Thermischer und Mechanischer Reize auf die Lorenzinischen Ampullen der Selachier, *Pflueger Arch. Ges. Physiol.* (Berlin), Vol. 263, p. 48 (1956).
28. Cauna, N., The Nature and Functions of the Papillary Ridges of the Digital Skin, *Anat. Rec.*, Vol. 119, p. 449, (1954).
29. Straile, W.E., Sensory Hair Follicles in Mammalian Skin: The Tylotrich Follicle, *Amer. J. Anat.*, Vol. 106, p. 133, (1960).
30. Bayliss, L., Principles of General Physiology, Longmans Green Co., Ltd., New York, (1960).
31. Cohen, M.J., The Function of Receptors in the Statocyst of the Lobster Homarus americanus, *Journal of Physiology*, Vol. 130, p. 9, (1955).
32. Lee, F. S., The Functions of the Ear and the Lateral Line in Fishes, *American Journal of Physiology*, Vol. 1, p. 126-144 (1898).
33. Pringle, J. W. S., On the Gyroscopic Mechanism of the Halteres of Diptera, *Phil. Trans. Roy. Soc., London*, Vol. 233, p. 347, (1948).



34. Davis, H., Initiation of Nerve Impulses in Cochlea and other Mechano-Receptors, In: "Physiological Triggers and Discontinuous Rate Processes", edited by T.H. Fullock, American Physiological Society, Washington, D.C. (1957).
35. Pumphrey, R.J., Hearing, Symposia Soc. Exp. Biol., Vol. 4, p. 1 (1950).
36. von Békésy, G., Experiments in Hearing, McGraw-Hill, New York, (1960).
37. von Békésy, G., Description of Some Mechanical Properties of the Organ of Corti, Journal Acoustical Soc. of America, Vol. 25, p. 770, (1953).
38. von Békésy, G., Rosenblith, W.A., The Mechanical Properties of the Ear, In: Handbook of Experimental Psychology, edited by S.S. Stevens, New York: Wiley, (1951).
39. Evans, H.M., A Contribution to the Anatomy and Physiology of the Air Bladder and Weberian Ossicles and Cyprinidae, Proc. Royal Society, London, Vol. 97, p. 545 (1925).
40. Walcott, C., van der Kloot, J.G., The Physiology of the Spider Vibration Receptor, Journal of Experimental Zoology, Vol. 141, p. 191 (1959).
41. Salpeter, M., Walcott, C., An Electron Microscope Study of a Vibration Receptor in the Spider, Experimental Neurology, Vol. 2, p. 232, (1960).



4. Photoreceptors

Helmholtz is reported to have said that if an optician made for him an optical system as poor as that of the human eye, he would not accept it. Whatever the specific geometrical limitations of the eye may be, the total visual apparatus is undoubtedly a wonderful mechanism.

This study has brought to notice some special features of the visual sense, which could lead to the development of several new devices.

Several of the remarkable properties of vision are associated with the visual acuity of the eye. Acuity is resolving power and can be defined in several ways according to the type of object being resolved. Normally, acuity is defined as the angular width of the smallest letter of an optician's eyetesting chart that can be reliably identified. With this definition normal human acuity is about 1 minute of arc ($1/3$ milli radian). However, this is the acuity at the center of the field of view, i.e. along the axis of regard. It is easily verifiable that acuity falls off very rapidly away from this axis. Thus if the eye is concentrated - i.e. fixated - on one letter of this text, which subtends 25 minutes at 12 inches, adjacent letters can be discerned only with difficulty, and those 4 or 5 places away are virtually unidentifiable. It is well known that in reading, the axis of regard is made to scan over the lines of text, so that the poor acuity away from the axis is no handicap. In fact it is an asset, since if we were able to resolve all the detail on a page without having to direct the axis of regard, the mind would be swamped with information. It is, of course, no accident that our reading and writing techniques are matched to the acuity characteristics of the eye and to the maximum data handling capabilities of the brain.

Although it is well known that our attention must scan a page of text in order to read it, it is perhaps not so widely appreciated that the scanning process is accomplished, essentially, by precise mechanical scanning of the optical axis of the eye over the page. The impression we get, however, is that the total visual impression is available all the time and that we merely direct our attention - a purely mental function - at each letter or group of letters. It is this illusion



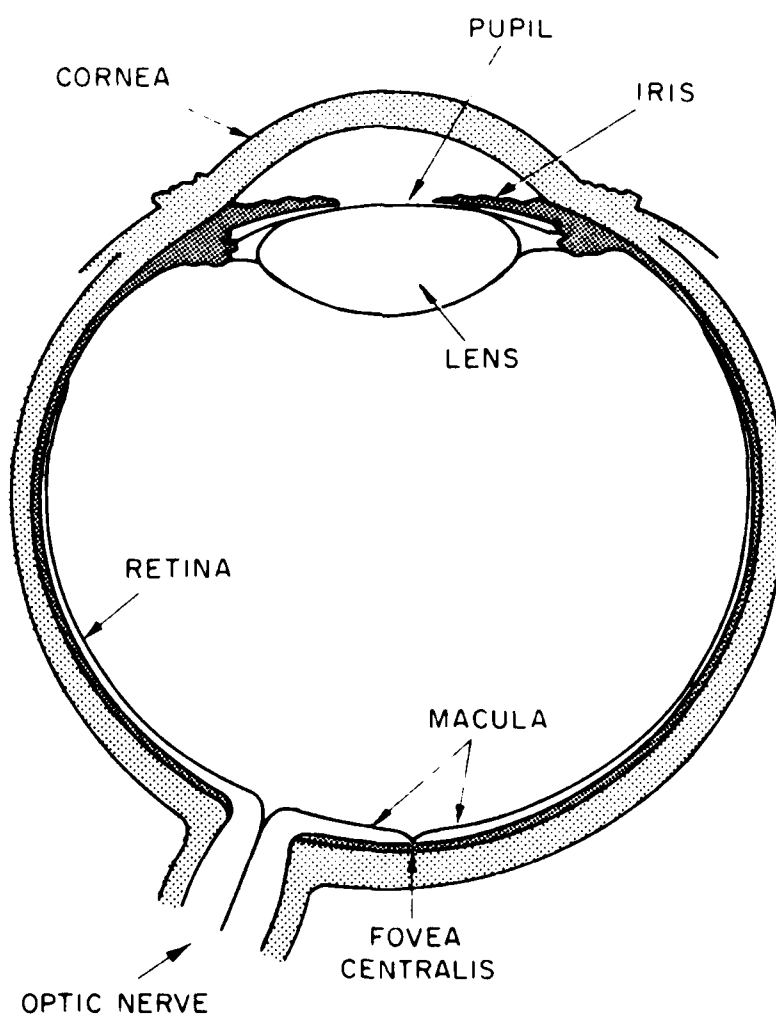
which allows us to believe that the eye presents the brain with a vast panoramic high definition picture. In fact, as will be shown later, the effective retinal image averages out to be equivalent to a relatively low definition picture.

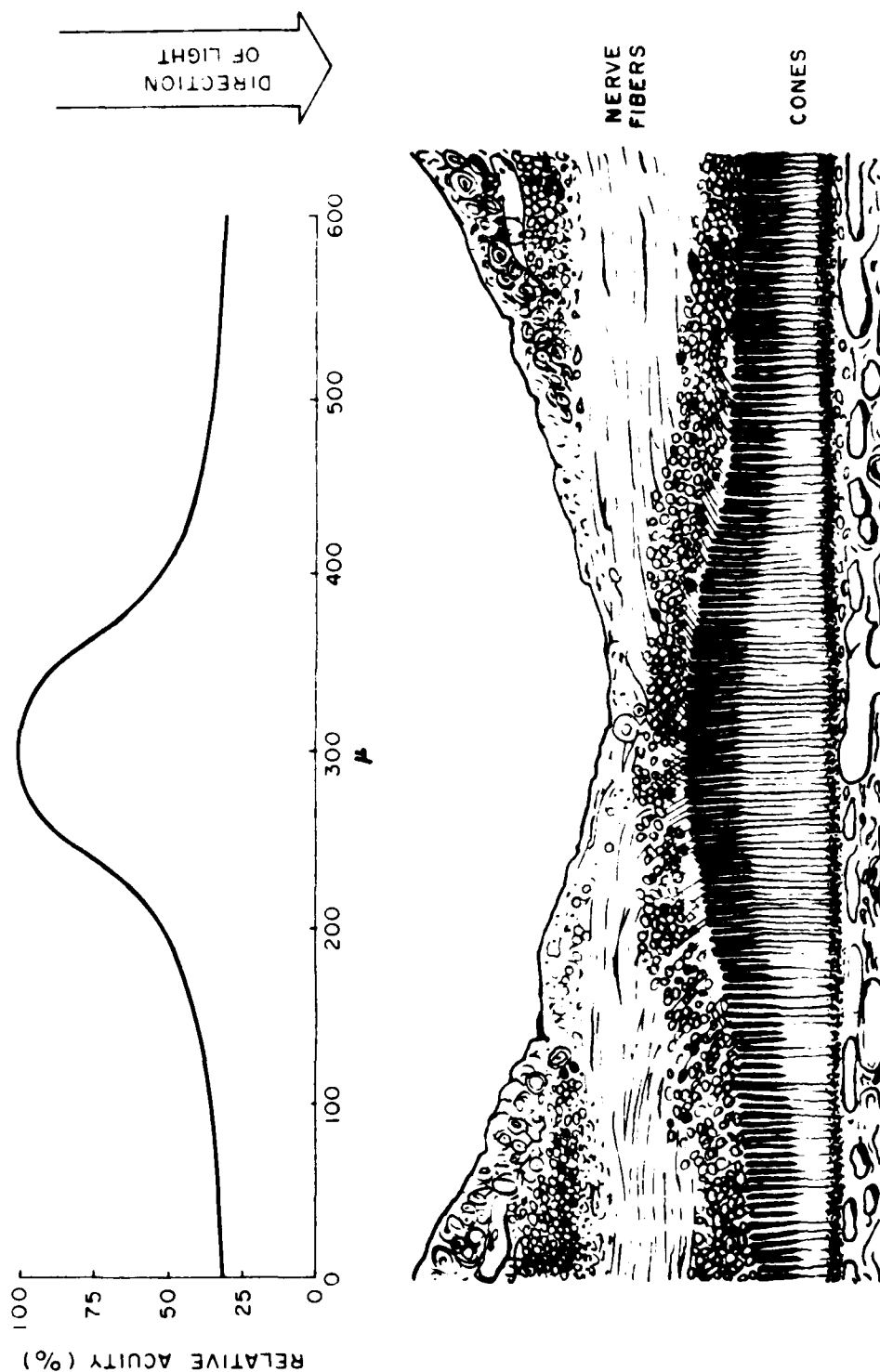
The acuity of the eye, as a function of angular displacement from the axis of regard, has its physical origin in the structure of the retina (Figs. 4-1 and 4-2). The photoreceptors - rods and cones - are located in the retina. In the central portion - the macula - most of the receptors are connected to the brain through individual nerve fibers. At the very center of the macula - the fovea - virtually all receptors have a "private line" to the brain. Elsewhere in the retina groups of photoreceptor cells are parallel connected to a single nerve fiber. The photoreceptors represent in some way a "grain" structure in the visual image, so that acuity must necessarily fall off away from the fovea.

Because acuity is so sharply concentrated at the fovea, the oculomotor apparatus that controls the orientation of the eyeball, relative to the head, is a most important part of the total visual mechanism. A great deal of spatial information is obtained directly from the oculomotor apparatus; we have the impression of knowing in great detail the physical structure of the room in which we are sitting, yet in any one position of the eyeball only a very small part of the visual scene can be observed with any sort of detail. Thus, large scale detailed spatial information cannot come from the retina alone. "Normally, information concerning the positions of objects in space is to a large extent dependent upon proprioceptor impulses arising in the ocular muscles. When, for example, we look at an object straight in front of us an image falls upon the macula of each retina. When we look at an object to one side the eyes are turned so that the images fall again upon precisely the same areas - the maculae. The actual position of the object - whether in front or to one side - is made known to us by afferent impulses set up in the muscles as they turn the eyes into position. Such impulses serve also as a basis for the nice correlation between visual sensations and various body movements. When, as a result of paralysis of certain muscles, the eye does not move with the sound eye, the impulses arising in the muscles of the latter convey the impression, nevertheless, that such a movement has taken place. Let us say the outward movement of the right eye is



THE HUMAN EYE





ALLIED RESEARCH ASSOCIATES, INC. (REF. 23)

LONGITUDINAL SECTION OF CENTRAL MACULA OF HUMAN EYE



paralyzed and the left, sound eye, is covered; when the latter turns to the right the patient believes the paralyzed eye does so to the same degree. The image in this eye, which continues to look forward, falls upon the nasal side of the retina. Since earliest experiences have taught him that an image falling upon the nasal part of the retina when this eye is rotated outwards represents an object well over to the right, he falsely project the image into this position, and makes the appropriate movements of the hand in an attempt to touch the object". (Ref. 1).

In addition to the conscious motion of the eye and motions in response to external stimuli, there are several small involuntary motions which are essential for vision. Recently ways have been found of stabilizing the image on the retina in spite of the involuntary motion of the eye and then, it is found, the visual conscious impression quickly fades out (Refs. 2 and 3).

In recent years very precise techniques have been developed for the measurement of eye motion - a subject referred to here as oculometry. These techniques have, so far, apparently only been applied to physiological and psychological research, the equipment involved has not been developed or engineered beyond the laboratory stage. There appear, however, to be numerous useful applications in other fields.

The reason that oculometry appears so fruitful a field of exploitation is that it allows a simple and effective "tap" to be made into one of the most important human sensory input channels. It is probably unique in this respect - i.e. connection to an input sensory channel will usually involve the surgical implantation of electrodes, and the data obtained may be subject to ambiguities arising from such effects as adaptation.

Whenever a man is made part of a machine system - i.e. as an aircraft pilot, a missile tracking camera operator, an anti-tank guided missile operator, etc. - he is almost always included in order that use may be made of his phenomenal power of pattern recognition. It is this human faculty that cannot be reproduced by a machine - and yet it is so often required. Thus it is only to be



expected that many useful applications of oculometry will be found at the man/machine interface. Oculometry will allow the optimum use to be made of the superior human faculties without the heretofore, inevitable limitations of the inferior human faculties - e.g. a man may often perceive a target yet be unable to move his hands quickly enough to bring a gun to bear on it. A machine can move and orient a gun much better than a man, but a man can pick out a target under conditions in which it would be impossible for a machine to do so.

Two general types of application of oculometry are discussed briefly below.

In command control systems, such as for example a wire guided anti-tank missile, a human operator is required to perceive the spatial relation of missile and target and then to generate guidance commands with a control stick. As already discussed, perception of spatial information stems largely from the oculomotor apparatus. Thus the technique of oculometry will allow this information to be extracted in the same form, and at the same time, as it is presented to the consciousness. In this way the sequence, in the normal use of a human operator, of translating the conscious perception of the spatial situation of missile and target into control stick motion, is eliminated. It is this part of the human operator activity that involves learning how to control the missile, and is responsible also for the noise, nonlinearity, inaccuracy and some of the sluggishness of the human operator. The oculomotor activity involved in both the normal and the proposed arrangement represents an irreducible minimum of activity for the human operator as it is essential for perception. It is, however, a very well practiced (over a lifetime) and highly developed neuromuscular activity. Thought of as a position servo, it has a very good position feedback loop to the brain via the retina, whereas the position feedback involved in control stick motion is the relatively crude sense of touch and pressure.

A second general application of oculometry would be to what might be called HI-FI TV.

In the early days of radio and the phonograph people heard music and voices reproduced artificially but nobody could then be deceived into mistaking a reproduced sound for the real thing. Nowadays, however, sound reproduction can be almost, if not entirely in some cases, indistinguishable from the real thing.



We are now able, by means of TV, to see pictures reproduced electronically. However, no one is ever deceived into mistaking a TV image for the real thing. Even if depth of contrast were improved, and the picture presented stereoscopically and in perfect color, still no one could mistake the stimulus obtained by sitting in front of the TV set with the real thing. The essential limitation of TV, as we know it, is field of view. A TV screen at normal viewing distance has a very small angular width. The limitation of such a narrow field can be readily appreciated by stopping down the eyes to an angle equivalent to a TV screen at 12 feet. From the point of view of gaining a total visual impression, it is almost like being blind when the eyes are restricted in this way.

Thus it is clear that to produce a HI-FI TV system, that would be virtually indistinguishable from reality, a very large angular field of view will be necessary. However, to increase the field of view of the domestic TV system to match that of the eye, and at the same time to maintain its grain structure quality, a bandwidth of over 2,000 mc would be required, compared to the 5 mc of normal domestic TV.

It is not necessary, however, to maintain the grain structure quality over the whole field of view. As has been pointed out earlier, the acuity of the eye falls off very rapidly away from the axis. Thus the grain structure of the TV system can be allowed to deteriorate away from the axis of regard without any noticeable effect. In this way the bandwidth requirements may be less, by orders of magnitude, than that even of the present domestic TV system.

In order to make a HI-FI TV system of the type discussed, two things are necessary:

- (1) a TV scanning system in which the grain structure is a function of the angular separation from the central axis of the system.
- (2) alignment of the TV camera to be under the sole control of the observer's oculomotor apparatus. Thus, this HI-FI system can only be used by one observer per camera.

It is shown later how the techniques of oculometry can be applied to make such a TV system possible.



It is emphasized that the design aim could be to produce a TV system that was quite literally indistinguishable from reality. The illusion would be complete: no visual impression would be gained of the TV monitor or its surroundings. The sole impression would be that derived by an eye located at the TV camera.

It is instructive to consider the implications of such a TV system. It is well known that our sensual contact with the world is almost entirely via the senses of sound and sight. The other senses do convey information but mainly for more or less closed loop functions such as stability, protection from damage, etc. Our ability to get the "feel" or "presence" of an environment depends almost entirely on the sense of sight and sound. Without them we are virtually isolated from the world, with them we maintain whatever contact our inner consciousness has with reality. The development of a truly HI-FI TV system would thus permit us to experience reality at a distance.

Useful applications of HI-FI TV can be envisaged for those situations in which it would prove easier or cheaper to transmit the HI-FI television pictures than to physically transport a human being. If sensual impressions identical to those obtained by being actually present can be derived by TV, there seems to be no logical reason (from the point of view of passive observation) for making the journey. An obvious example, of course, is travel to the moon and planets for the purposes of exploration. In this case, enormous advantages accrue by sending expendable TV equipment rather than a non-expendable astronaut on such journeys. However, the delay involved in the round trip communications link - due to the finite velocity of electromagnetic waves - would probably restrict the useful range between camera and monitor, to about 10,000 miles for a truly HI-FI TV system. Beyond this range an appreciable delay would occur between the observer moving his eye and the reception at the monitor of the newly oriented picture. Between earth and moon the delay is about 2 1/2 sec so there can be no suggestion that remote TV over this distance could duplicate the visual sense.

The 2 1/2 second transmission delay to the moon and back is a disadvantage for a conventional remote TV System, as discussed in Reference 4. However, by



accepting this limitation, some form of moon exploration can be undertaken almost at once. It is shown later that, by using an oculometrically controlled TV System, a very substantial saving in bandwidth can be effected. This is important as power requirements are directly related to bandwidth.

4.1 The Human Eye

From the systems point of view the human visual apparatus may be considered as comprising: an optical system of which the aperture, focal length, and axis of pointing are controlled by nervous and muscular action - a photo sensitive screen, i.e. the retina, on which an inverted optical image is formed of the scene being viewed - an information transmission system, i.e. the nerve fibers linking the retina to the brain - and an interpretive mechanism whereby the data transmitted by the photosensitive elements of the retina is processed to yield useful information for the consciousness. This interpretive mechanism is not well understood; it is thought to occur in the retina itself to some extent, as well as in the brain.

4.1.1 The Retina

The retina is the screen at the back of the eye on which is formed a real image of the scene being viewed. It consists of a number of layers of nervous elements. As light passes into the retina from the vitreous humor it meets first an outer layer of nerve fibers that connect the photoreceptors to the brain via the optic nerve. Near the back, or base, of the retina is the layer of rods and cones which are the actual photoreceptive elements. These are connected to the nerve fibers at the upper layer of the retina by various types of neural interconnections. In the central foveal region of the retina each photoreceptor has its own nerve fiber line to the brain. At other points on the retina groups of photoreceptors must share a nerve fiber, and thus visual acuity is reduced. At the center of the macula- the fovea -the upper retinal layer of nerve fibers thins out, and the rods and cones are thinner and more closely packed as shown in Figure 4-2.



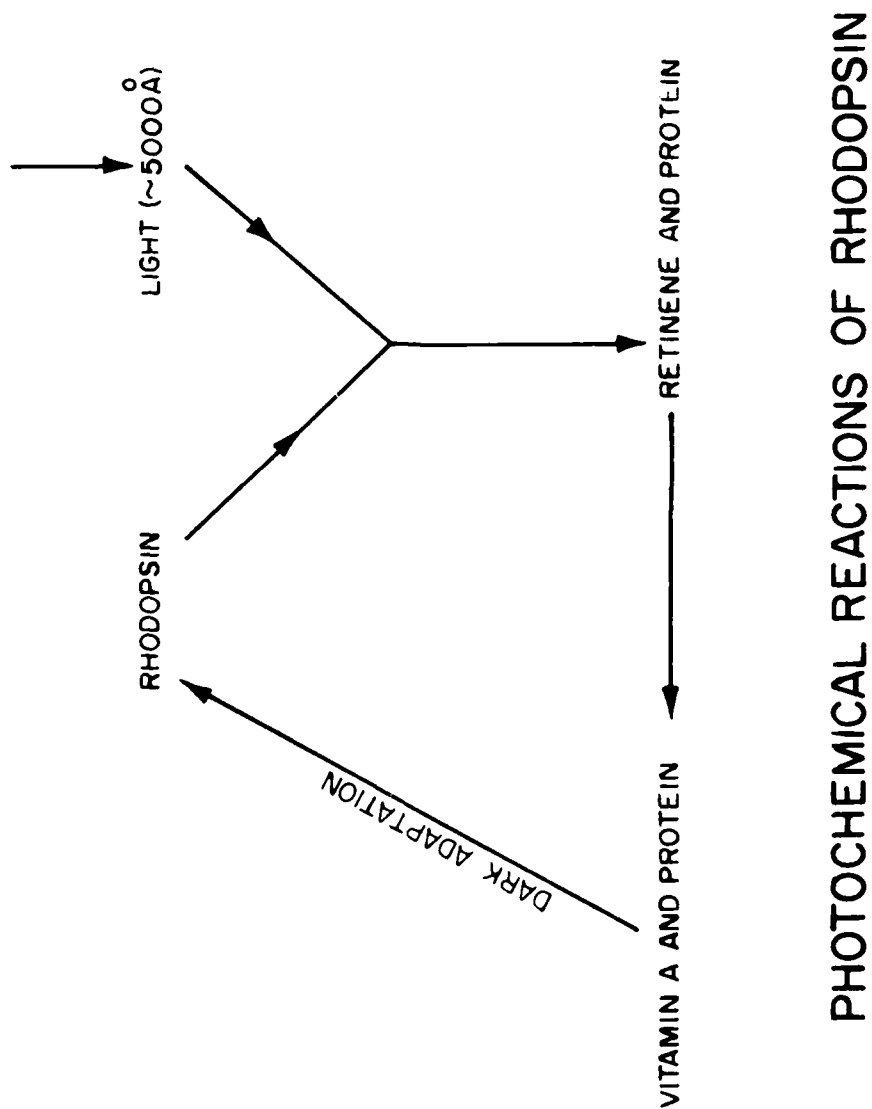
The photoreceptors respond by means of a photochemical reaction involving the pigments rhodopsin, for the rods, and iodopsin for the cones.

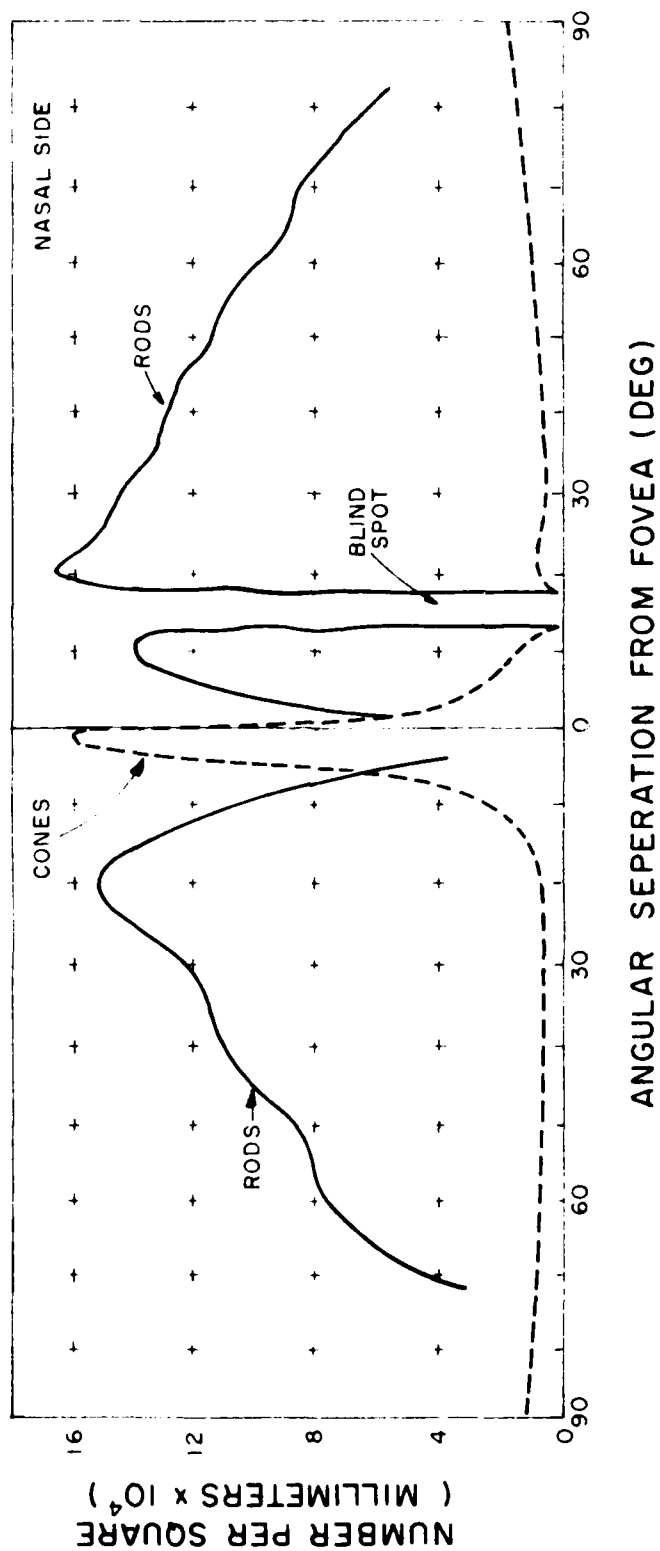
The initial step in the visual stimulation of the optic nerves is the photochemical reaction illustrated in Figure 4.3. The final regeneration of rhodopsin from vitamin A takes place in the living retina during the dark adaptation period. Stimulation of the optic nerve is, presumably, effected either directly or indirectly by the presence at the retinal synapses of the breakdown substances of rhodopsin.

The signals occurring in the optic nerve fibers in response to a visual stimulus consist of a series of short pulses, the pulse repetition rate being proportional to the intensity of the visual stimulus. An action potential can be detected between the retina and cornea, the amplitude of which is related to the intensity of the visual stimulus. The cones, which are about 28-85 μ long, are concentrated in the central part of the retina. In the fovea there are virtually no rods. At the periphery of the retina the receptors are virtually all rods which are about 40-60 μ long.

The distribution of rods and cones in the retina is illustrated in Fig. 4-4. According to most texts, the cones function only at relatively high light intensities, and the rods cease to respond above a certain light intensity. At very low levels of illumination only the rods function. There is a range of intensity over which both rods and cones function. The light adapted cone vision is called scotopic and the dark adapted rod vision photopic. The range of intensities over which the eye will operate is very large- 10^5 - and this is explained, according to the duplicity theory above, partly by the existence of two types of photoreceptors with different operating ranges, and partly by the large dynamic range of each type.

In another theory (Ref. 5) the rods and cones are considered to be more or less equally responsive. Calculations quoted in Reference 5 show that there are not enough quanta available to stimulate each individual receptor, except at the highest levels of illumination. Thus since cones are chiefly connected singly, or in small groups, to a nerve fiber, very few cone connected nerve fibers become active in low or moderate light intensity. The rods, however, are mostly connected in large groups to a nerve fiber so that rod connected fibers will





DISTRIBUTION OF RODS AND CONES IN THE HUMAN RETINA
(ADAPTED FROM REF 24)



continue to respond even though the number of quanta per rod may be very low. At the threshold of vision the quantum flux is equal to one quantum per photoreceptor per 10^3 secs. At an intensity 10^5 times this the number is 10^2 per photoreceptor per second. The summation time of the photoreceptors may be taken as 0.1 sec, so that even at this high level of illumination only 10 quanta per summation time arrive, on the average, at each photoreceptor. These figures do not allow for losses in transmission through the eye or in the photoreceptors themselves. Thus it is clear that for most of the range of illumination over which the eye operates the actual origin of the visual stimulus is a photochemical reaction involving only a very few quanta.

This theory explains the relationship between acuity and illumination, in effect, by asserting that human visual acuity approaches the limit set by the quantum nature of light energy, the size of the pupil of the eye, and the summation time.

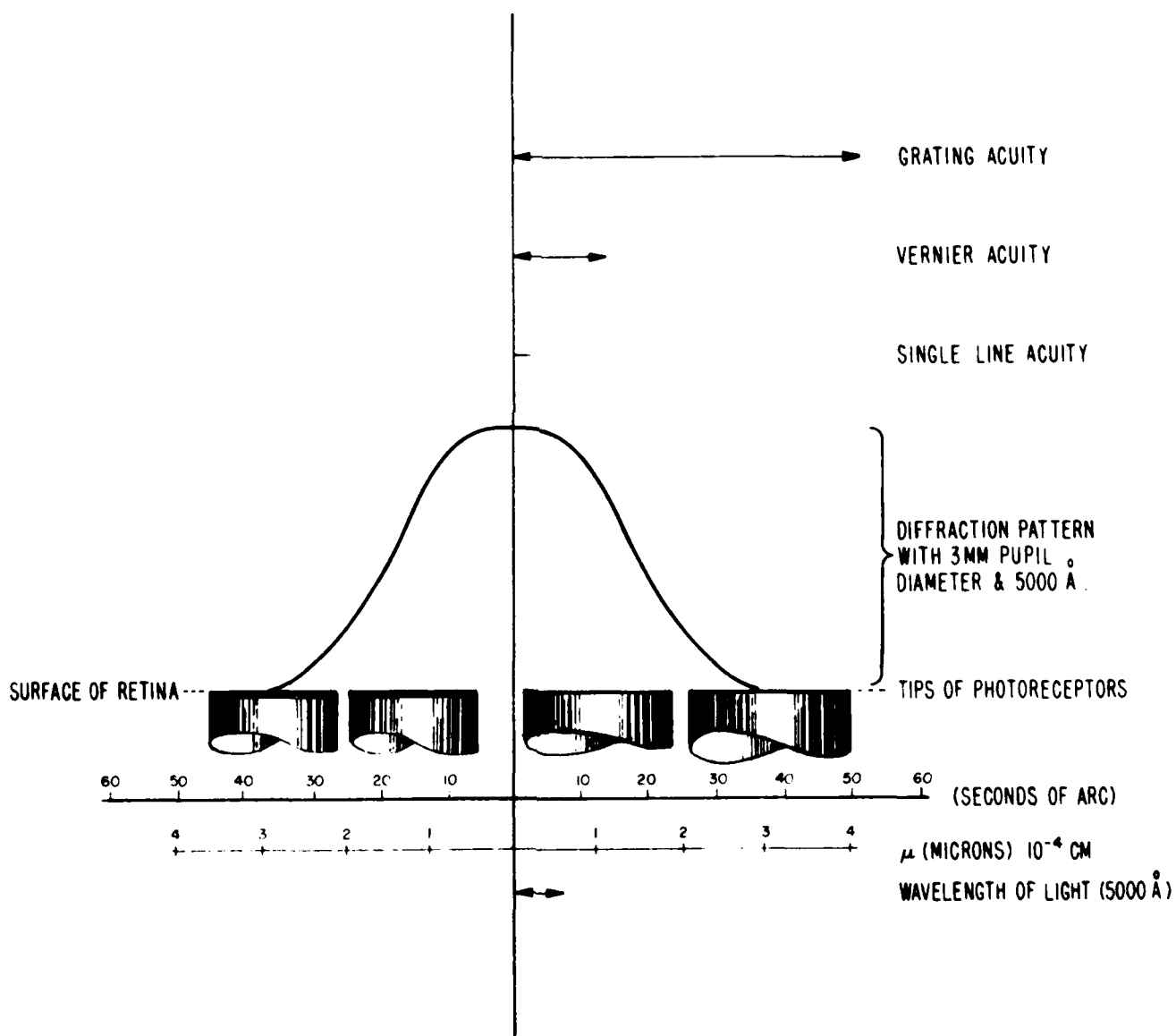
In both theories the visual action changes from cones to rods as illumination is decreased, thus explaining color blindness at low illumination.

Many theories of color vision have been advanced. However, it is known that it is the cones that are color responsive. A recent theory (Ref. 6) has suggested that the individual cones act as optical wave guides to produce standing waves, the spatial pattern of which is a function of the color. Thus, color discrimination could be due to a spatial pattern in the nervous stimulation mechanism within each cone, and not to the presence of several different types of photochemical material.

The acuity of the human eye depends to some extent on the type of object being resolved. In Figure 4-5, the grating, vernier and line acuities are shown, together with the diffraction pattern of a point source, and an indication of the size of the individual photoreceptors. It can be seen that the size of the photoreceptors is reasonably related to the diffraction pattern, and that the grating acuity corresponds to both. However, the single line acuity is considerably better than might be expected. It was thought at one time that the small involuntary motion of the eye was the mechanism by which this fine detail was resolved by the eye. However, experiments have been reported (Ref. 7) which tend to disprove this theory.



FOVEAL ACUITY





It has been pointed out that the considerable length of the cones poses a difficulty in understanding visual acuity since a point image - actually the diffraction pattern of a point i.e. a circle about 5μ in diameter - may be produced by a cone of light up to approximately 20° angular width. Thus after traversing the length of a cone - a distance of about 60μ - it would be expected to cover a circle of diameter of about 20μ (Fig. 4-6). A tunneling theory has been proposed to explain why this does not happen. It is postulated that the refractive indices of the cone and the medium in which it is situated are such that the 20° cone of light is captured and confined to the cone by total reflection at the walls. The theory involves the tapering of the cones, and is able to explain the Stiles and Crawford effect namely the reduction in the sensitivity of cones to light entering the pupil near its edge, compared to that entering in the center of the pupil (Ref. 8).

The highest acuity of the eye is concentrated very sharply at the fovea, as indicated in Figure 4-2. It has been claimed that a drop in acuity can be detected only 3.5 minutes of arc from the center.

There are estimated to be 120 million rods and 6 million cones in the retina and one million nerve fibers in the optic nerve.

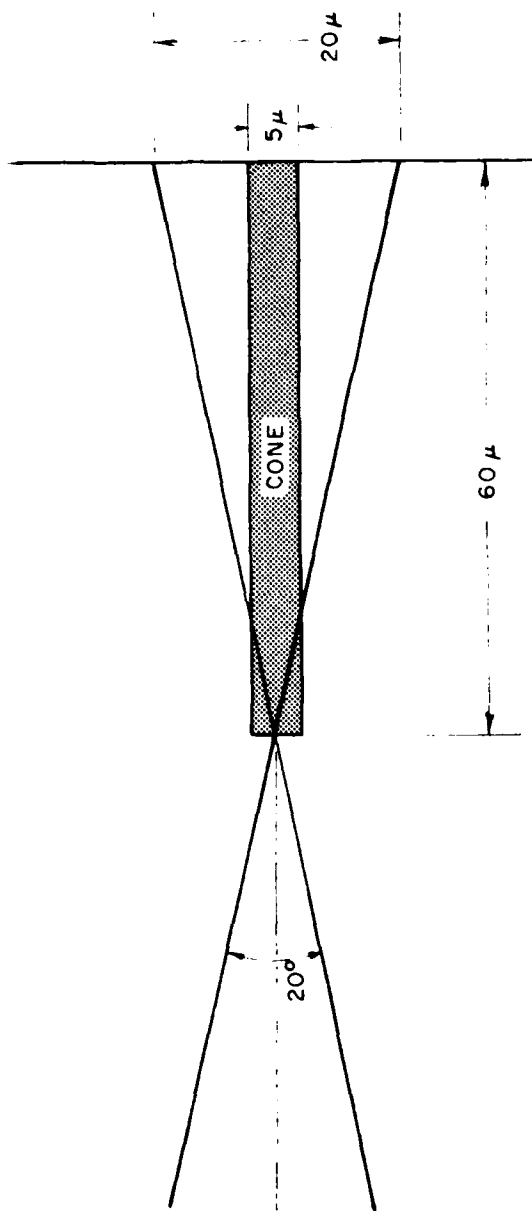
4.1.2 Optical System

The essential features of the optical system of the human eye are shown in Figure 4.7.

4.1.3 Oculomotor Apparatus

The acuity in human vision is concentrated very sharply at the center of the eye. Little published data on peripheral acuity has been found. In Figure 4.8 the data taken from Reference 8 giving the acuity out to 1° has been extended by some approximate subjective measurements. The peripheral acuity is so poor that our visual sense depends very considerably on the oculomotor apparatus.

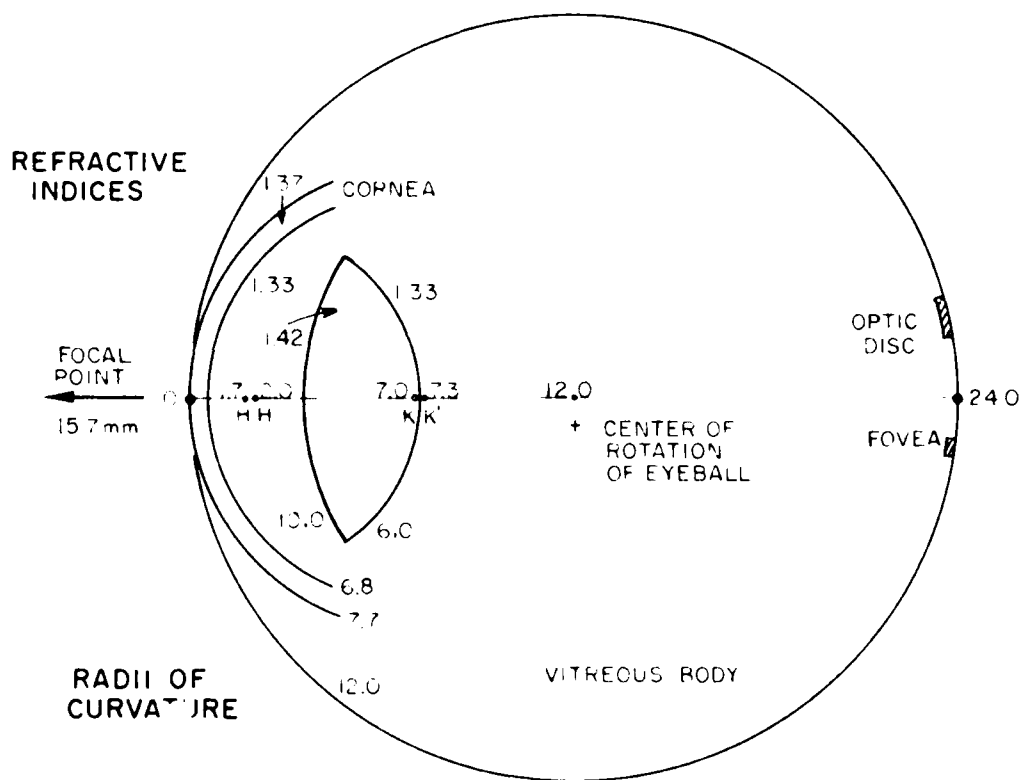
The eye is rotated in its socket by muscular action in response to commands from the conscious and subconscious centers of the brain. In addition to reflex responses there are three involuntary motions of the eyeball which occur when the eye is steadily fixated on a test point (Refs. 9 through 13):



ILLUMINATION OF A SINGLE RETINAL CONE

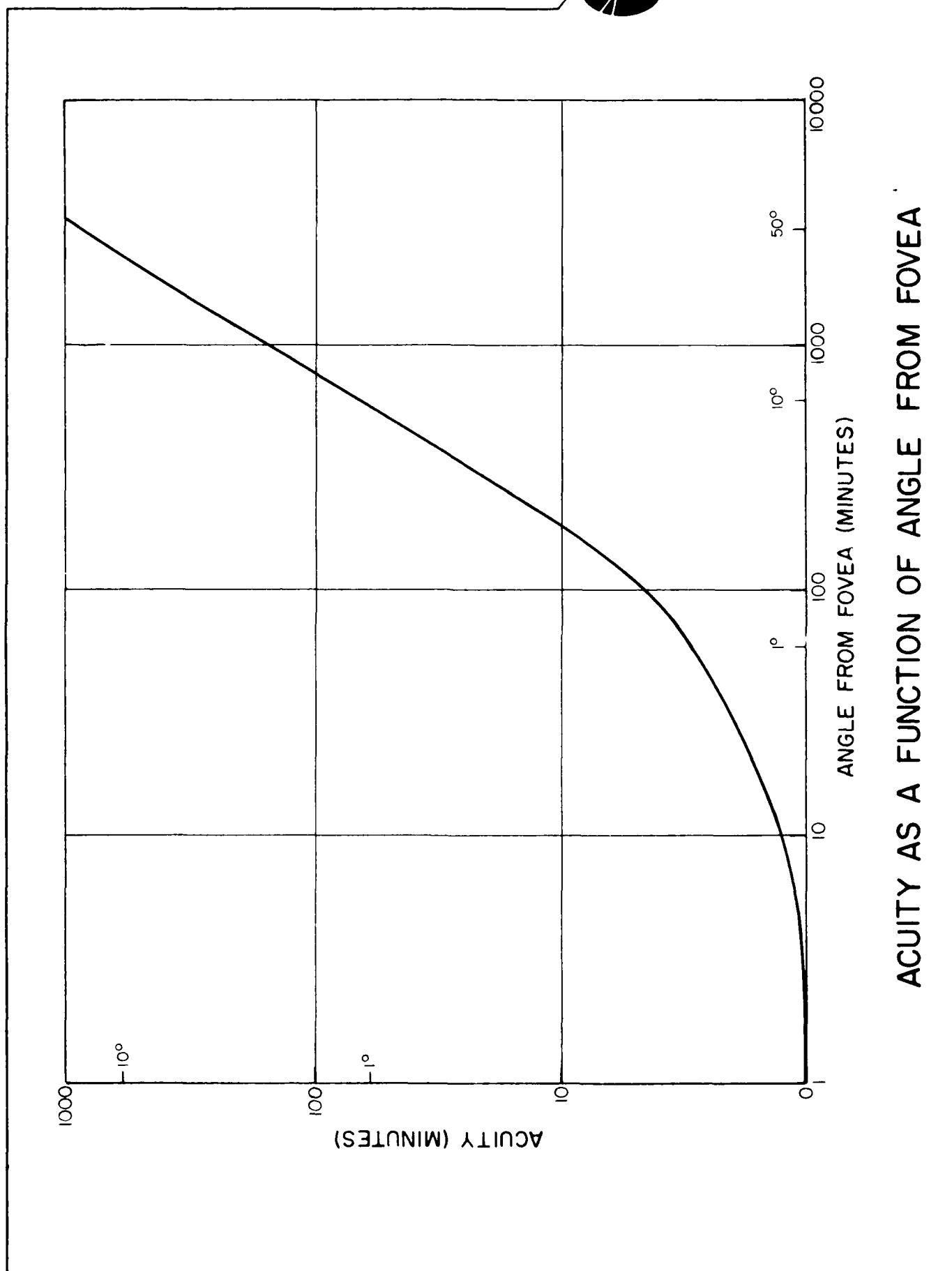


SCALE DRAWING OF OPTICAL SYSTEM OF THE HUMAN EYE



ALL DIMENSIONS IN MILLIMETERS

K K' = NODAL POINTS
H H' = PRINCIPAL POINTS





1. A very small continuous tremor at 40-100 c.p.s. of amplitude between 10 and 20 seconds of arc.
2. A slow steady drift of up to about 5 minutes of arc in 0.1 to 0.3 seconds.
3. A sharp flick or saccade of about 5 minutes of arc every 0.1 to 0.3 seconds.

During a steady fixation the total integrated motion of the eye is hardly ever more than about 10 minutes of arc.

These involuntary eye motions during steady fixation have been the subject of considerable research. One of the most remarkable discoveries has been that they are essential for the continued generation of a conscious visual stimulus. Several methods (Refs. 2 and 3) have been found of stabilizing the image on the retina so that although the eye may move, the image remains stationary. It is found that under these conditions the visual conscious sensation rapidly fades out. In one series of experiments (Ref. 3) the converse of stabilizing the retinal image i.e. arranging for it to move twice as much as in normal vision, was undertaken. It was found that a marked improvement was obtained in the acuity for periods of fixation of 1 sec. or more. Moreover, under these conditions of viewing "the impression was that the target was "locked in place" so that steady fixation was effortless, automatic. It may be noted that in this condition the usual cues for fixation are exaggerated. The eye muscles are provided with double the normal feedback from any drifting of the image away from optimal fixation" (Ref. 3). For periods of very short fixation, less than 0.2 sec, the best acuity is obtained with the retinal image stabilized. In any viewing condition it takes about 1/2-1 sec for maximum acuity to develop.

It is possible to experience something of the effect of the fade out of the visual field due to a stabilized retinal image without any apparatus. If a bright point, such as a star, is fixated under conditions of very low illumination of a smoothly textured background, it will be found that the visual field begins to fade in a second or two, but comes back suddenly, apparently as fixation is lost.



It has not been established whether the involuntary motion of the eye serves any purpose other than to prevent this fade out. It was believed at one time that the high acuity of the eye relative to the cone size, was due in some way to this motion. Only one report (Ref. 10) has been found of research into involuntary eye motion in animals: in this case it was a cat, in which such motion was observed.

There is evidence (Ref. 14) that the small tremor and slow drift of the eye are effects due to a slight instability in the neuro muscular feedback system of the oculomotor apparatus. The saccades, on the other hand, are apparently triggered by the magnitude of the displacement of the retinal image due to the other two motions; i.e. it is an unconscious attempt to improve acuity.

In binocular vision both eyes follow each other, in a fixation, to an accuracy of about 10 minutes of arc, which is comparable to the individual involuntary motions of each eye. The slow drift and tremor in the two eyes are uncorrelated. The overall correlation is achieved by the saccadic motion which is strongly correlated between the two eyes. The maintenance of binocular fixation does not seem to be dependent on a direct response to, or sensing of, vergence error. Rather it appears to be dependent on the saccadic responses of the two eyes to their own fixation errors (Ref. 11). It has been shown that involuntary eye motions are not important for the perception of depth in stereoscopic vision (Ref. 15).

The saccadic flick motion of the eye initiated consciously to change the fixation point, has a latency of 0.2 second and a maximum speed of up to $500^{\circ}/\text{sec.}$ for large movements and $100\text{-}200^{\circ}/\text{sec.}$ for small movements. In smooth pursuit motion the latency is again 0.2 sec. The maximum speed of smooth pursuit is between $30\text{ and }60^{\circ}/\text{sec.}$ A reduction in acuity is observed in viewing moving objects: at a speed of $100^{\circ}/\text{sec.}$ it is halved. It is postulated that this is caused by a reduction of intensity gradients due to inadequate tracking by the eye. In general, acuity in observing moving objects can be improved by using higher levels of illumination (Refs. 9, 16).



4.1.4 The Information Content of the Retinal Image

The information content of a picture can be taken as the product of the number of "grain" elements of which it is composed and the number of bits of information per grain element. The latter quantity is related to the effective overall signal/noise ratio and is apparent as the depth of contrast in the picture.

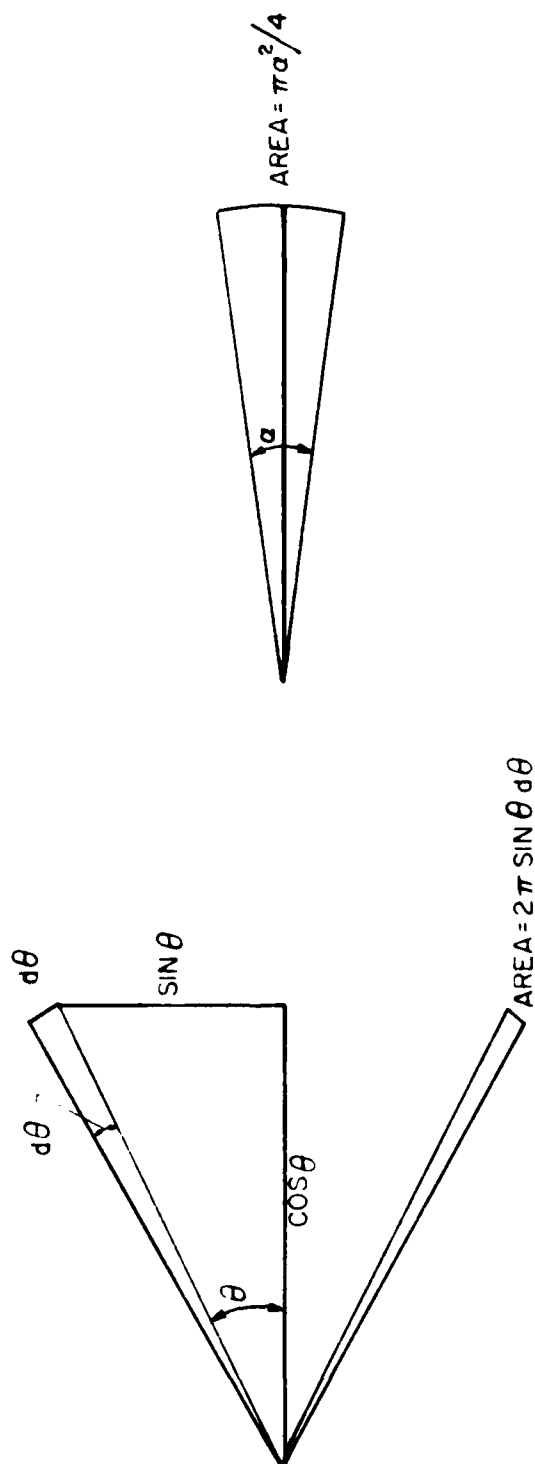
The grain structure of the retinal picture is obviously dependent on the acuity of the eye. As discussed previously this is not uniform over the whole field of view but is concentrated sharply along the axis of regard and falls off towards the periphery of the visual field. Little published data on peripheral acuity has been found. In the course of this study, some approximate measurements have been made of human acuity as a function of angular displacement from the axis of regard. The results are plotted in Figure 4-8 which also contains data taken from Reference 7 on the acuity/angular-displacement relationship within 1° of the axis of regard. As plotted in Figure 4-8, acuity has been taken as the angular width of a just discernible test letter.

An empirical functional relationship has been determined between acuity (α , in minutes) and angular displacement (θ , in minutes) from the axis of regard:

$$\alpha \approx 1 + 0.0353 \theta \quad 0 < \theta < 100$$

$$\alpha \approx 10^{-5.44} \theta^{3.22} \quad 100 < \theta$$

An approximate value for N , the number of separately resolvable grain elements in the entire visual field can now be deduced. The area of a sphere of unit radius that lies between θ and $\theta + d\theta$ of the central axis is $2\pi \sin \theta d\theta$ (Fig. 4-9). The area of the least resolvable detail is $\frac{\pi \alpha^2}{4}$, thus dN , the number of grain elements between θ and $\theta + d\theta$ from the central axis, is given by



CALCULATION OF N



$$dN \sim \frac{2\pi \sin \theta d\theta}{\pi \frac{a^2}{4}} \sim \frac{8\theta d\theta}{a^2}$$

Thus

$$N = 8 \int \frac{\theta d\theta}{a^2}$$

where suitable limits must be given to the integral.

The number, N_p , of grain elements in peripheral vision further than 100 minutes of arc from the axis of regard is given by;

$$\begin{aligned} N_p &\approx 8 \int_{10^2}^{6 \cdot 10^3} 10^{5.44} \theta^{-2.22} d\theta \\ &= \frac{-8}{1.22} 10^{5.44} \left[\theta^{-1.22} \right]_{10^2}^{6 \cdot 10^3} \\ &\approx \frac{8}{1.22} 10^{5.44} 10^{-2.44} \\ &= \frac{8}{1.22} 10^3 = 6,500 \end{aligned}$$

The number N_f of grain elements in foveal vision within 100 minutes of arc of the axis of regard is given by

$$N_f = 8 \int_0^{100} \frac{\theta d\theta}{(1 + 0.035\theta)^2}$$



$$\begin{aligned}
 &= \frac{8}{(0.035)^2} \left[\int_0^{100} \frac{(1 + 0.035 \theta) d \theta 0.035}{(1 + 0.035 \theta)^2} - \int_0^{100} \frac{0.035 d \theta}{(1 + 0.035 \theta)^2} \right] \\
 &= 6,540 \left[\log_e (1 + 0.035 \theta) + \frac{1}{(1 + 0.035 \theta)} \right]_0^{150} \\
 &= 6,540 \left[\log_e 4.5 + 1/4.5 - 1 \right] \\
 &= 6,540 \cdot \left[0.25 \right] \\
 &= 1,830
 \end{aligned}$$

The total number, N_t , of resolvable elements in the visual field is thus

$$N_t = N_f + N_p = 8300 \sim 10^4$$

These calculations are approximate and intended only to give an estimate of the order of magnitude involved. Since the eye can actually distinguish a test letter when it subtends an angle equal to the acuity, as here defined, there may be perhaps 10 elementary grain elements in what has been taken as the smallest resolvable area, $\frac{\pi a^2}{4}$. Thus the total number of elementary grain elements in the retinal picture is probably of the order of 10^5 .

Assuming that the retinal image is transmitted to the brain directly, without significant breakdown or processing by the retina, this figure of 10^5 effective grain elements agrees reasonably well with the number of nerve fibers, $\sim 10^6$ - in the optic nerve. Several factors can be postulated as being responsible for the 10:1 ratio of nerve fibers to grain elements.

It is conceivable that there is more than one fiber per unit element, here defined, - possibly because the peripheral acuity may be limited by optical defects rather than grain structure, and also the method of transmitting color



information to the brain from each cone may involve multiple nerve fibers rather than a complex modulation process- as in domestic color TV - within a single fiber. It is also possible that the peripheral acuity has atrophied, due to lack of use, and could be improved with practice.

Allowing 4 bits of information per elementary grain element (i.e. 16 brightness levels) and a picture repetition rate of 5 per second, it can be seen that the information flux at the retina is probably of the order of 2 million bits per second. The information flux to the consciousness is, of course, very much less than this, suggesting the existence of some form of mental scanning of the data available at the retina, similar to the mechanical scanning by the eyeball.

It is very likely that considerable data processing of the retinal image is performed by the retina itself. In this way the information flux along the optic nerve may be less than the figure of 2 million bits / second indicated above.

4.2 Animal Eyes

4.2.1 The Frog

The frog's eyes move to compensate for body or head motion, but not for seeking out points of interest in the visual field as does the human eye. The retina does not have a central area of high acuity.

The frog does not appear to see stationary detail. It is said that he will starve to death if surrounded by stationary food. Recent investigations (Ref. 17) have demonstrated that the photographic image as detected by the rods and cones is subject to logical operations in the retina before transmission to the brain. The neural sequence is: photodetection by rod or cone - complex synaptic connections (within the retina) to the ganglion cells - transmission over the optic nerve fibers to the brain.

There is a synaptic path from each ganglion cell to many rods and cones and vice versa. These connections are such that each ganglion cell responds to a particular pattern in a part of the retinal image field. There is extensive overlapping of the retinal areas that activate each ganglion cell.

Five classes of ganglion response have been described. These are, in terms of the retinal pattern-



1. sustained edge detection
2. convex edge detection
3. changing contrast detection
4. dimming detection
5. darkness detection

The five classes of ganglion cells are uniformly distributed across the eighth layer in the retina.

It is apparent that the function of the retina in the frog is not to transmit information about the point to point pattern of distribution of light and dark in the image formed on it. Its function is rather to analyze the image at every point in terms of a few qualitative contexts (standing edges, curvatures, changing contrast, and local lessening of light intensity).

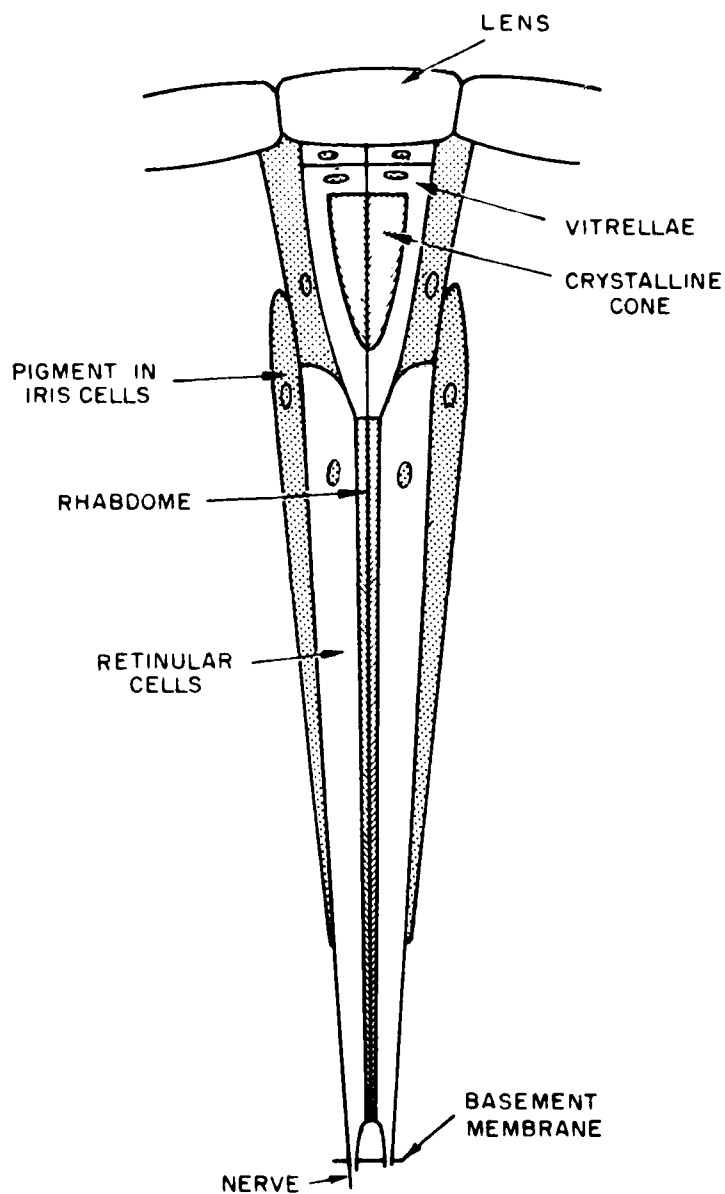
4.2.2 Arthropod Eyes

The eyes of most arthropods consist of one or more units - ommatidia - each of which each is a vertical bundle of cells over which is formed a lens. An ommatidium is illustrated in Figure 4-10. A true compound eye is composed of a number of ommatidia - each one with its own lens system. In most compound eyes the pigment flows up and down the ommatidia according to the light intensity. In dim light it is retracted to leave the sides of the ommatidium exposed, in bright light it extends to optically separate the whole ommatidium.

Vision takes place in two ways according to the situation of the pigment. In bright light the pigment ensures that a narrow pencil of almost parallel rays falls on the retinula. There is formed in this mosaic vision, an apposition image, composed of as many points of light as there are ommatidia. In dim light when the pigment is retracted, each ommatidia throws a complete image of the greater part of the field of vision onto the retinal layer and the images together form a supposition image, falling in such a way that corresponding points are superimposed. Supposition images are less sharp than apposition images but are formed with less loss of light.



OMMATIDIUM





4.3 Applications

4.3.1 Photodetection

4.3.1.1 Sensitivity of the Human Eye

Consider the eye viewing a screen which has a brightness of one lambert. Assuming a reflection factor of unity, one square meter of the screen radiates one lumen. Thus an element of detail of area A of the screen will give rise to a light intensity of $\frac{A}{\pi} \cos \theta$ candles, in a direction making an angle θ with the normal to the screen. Let the eye be placed a distance r from the screen and directly over this element A of detail, so that the detail subtends an angle α at the eye (see Figure 4-11).

$$\therefore A = \frac{\pi}{4} (r \alpha)^2$$

Let d be the pupil diameter of the eye; then the light flux entering the eye is

$$\left(\frac{\pi d^2}{4} \cdot \frac{1}{2\pi r^2} \right) \frac{\pi r^2 \alpha^2}{\pi 4} = \frac{d^2 \alpha^2}{32} \text{ lumens}$$

Assume $d = 3 \text{ mm} = 3 \times 10^{-3} \text{ m}$

$\alpha = 2 \text{ minutes} = 1/1800 \text{ radian}$

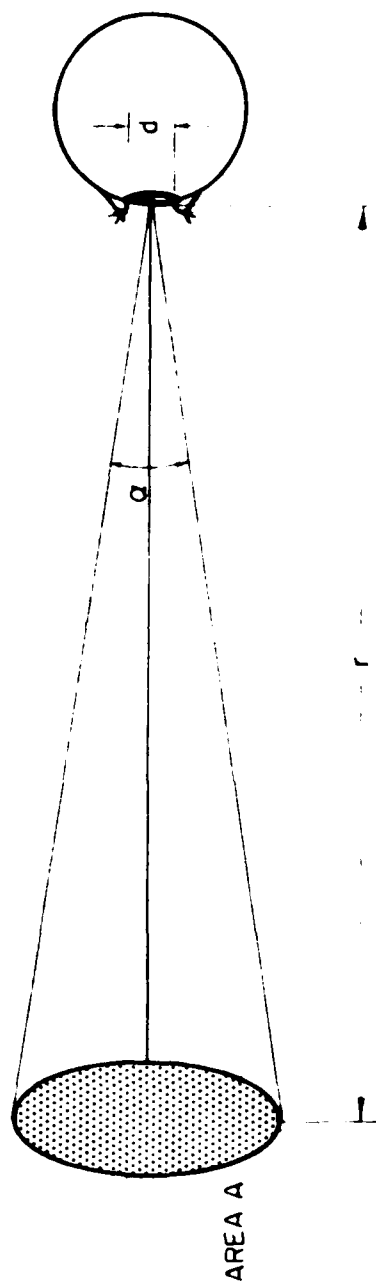
Therefore,

$$\text{flux} = \frac{9 \times 10^{-6}}{32 \times 1.8^2 \times 10^6} = 8.7 \times 10^{-14} \text{ lumens}$$

At a wavelength of $5600 \text{ \AA}$ a flux of 2.42×10^{-16} lumens is equivalent to a flow of 1 quantum per second.

Therefore, at a brightness of 1 lambert, detail of 2 minutes of arc angular width irradiates the eye with a flux of 3.6×10^2 quanta/second. The summation time of the eye is approximately 1/10 of a second, thus the basic physical stimulus at this resolution is caused by a packet of only 36 quanta. (It may be noted that this calculation agrees well with the results previously quoted from Reference 5).

The brightness level chosen for this calculation - i.e., 1 lambert - is about in the middle of the human visual range. The outdoor illumination in sunlight is of the order of 100 lumens/square meter (giving a brightness of 100 lamberts for



ILLUMINATION AT EYE



in ideal white diffusing surface). Desk work usually requires about one to two lumens per square meter. The threshold of vision is at about 10^{-3} lumens/square meter.

The optical efficiency of the eye is fairly poor (1% - 10%) - i.e., most quanta incident at the eye do not succeed in triggering the photochemical neural response. It is clear therefore that the rods and cones must be effectively stimulated by only a very few, possibly individual, quanta.

4.3.1.2 Photomultiplier Tube Sensitivity

The photomultiplier tube can be taken as an example of a very sensitive (inanimate) detector of visible radiation.

It consists of a photoemissive cathode situated near a number of multiplying anodes. The purpose of these anodes is to multiply, by the mechanism of secondary emission, the flow of electrons from the cathode. The multiplication factor that is achieved in commercial tubes is of the order of 10^6 .

As in any device, the effective sensitivity of a photomultiplier tube is limited by noise. If the phototube had no built-in amplification (i.e., electron multiplication), but was instead coupled directly to a vacuum tube amplifier, the limiting noise would be either the thermal noise of the input impedance of the amplifier (i.e., the load impedance of the phototube) or else the shot noise in the first stage of the amplifier. However, by using electron multiplication in the phototube itself, the phototube output signal level can be raised well above the inherent noise of any succeeding amplifying stages. With this arrangement the limiting noise becomes the shot noise associated with the dark current of the phototube. The dark current is due principally to the thermionic emission of electrons from the photocathode that occurs even when no light falls upon it. (Shot noise is the statistical variation of an electric current arising from its discrete - rather than continuous - structure of individual electronic charges).

The characteristics of the Dumont type 6467 photomultiplier tube are given below, in abridged form, as an example of a commercially available photomultiplier tube.



Spectral Response	3250 Å - 6125 Å
Cathode luminous sensitivity	60 u A/lumen
Anode luminous sensitivity	12 A/lumen
Anode dark current	.05 uA
Current amplification	215,000
Window diameter	1 inch

The cathode dark current is $.05 \div 215,000 \mu A$, i.e., $2.33 \times 10^{-13} A$. Thus the dark current shot noise is equivalent to a cathode current of:

$$i_{rms} = \sqrt{2ei} = \sqrt{2 \times 1.6 \times 10^{-19} \times 2.33 \times 10^{-13}}$$

$$= 2.72 \times 10^{-16} \text{ amps for one cycle bandwidth}$$

This corresponds to a cathode illumination of $2.72 \times 10^{-16} \times \frac{10^6}{60}$ lumens, i.e., 4.55×10^{-12} lumens for one cycle bandwidth. Over a 10 cps bandwidth (corresponding to the 1/10 second summation time of the human eye) this figure would be 1.4×10^{-11} lumens. It represents the limit of detectability since a light signal below this value would be swamped by the shot noise. Thus this phototube is about 160 times less sensitive than the elementary photoreceptors of the retina.

If a photomultiplier tube is cooled (say to $70^{\circ}K$) the thermionic emission from the cathode can be reduced by a factor of 10^6 . This would reduce the dark current shot noise by a factor of 10^3 and thus increase the effective sensitivity of the tube to 1.44×10^{-14} lumens over a 10 cps bandwidth. This is about 6 times the sensitivity of the eye. However, at this extreme sensitivity - virtually at the quantum limit - other effects, not here considered, may become significant.

4.3.1.3 The Significance of Size

The very small size of the rods and cones allows the resolution of a large amount of detail by an eye that is small, light, and requires very little power to operate it or to move it in its socket. It is postulated that the apparent invulnerability of the eye to thermal noise - that limits the performance of a photomultiplier tube at room temperature - is also attributable to the very small size of the individual rods and cones. In a photomultiplier cathode there are a large number molecules and electrons (all subject to thermal disturbance) and yet when the tube



is operating at very low light levels there are only a few quanta available for stimulating photoemission. Thus it is to be expected that the ratio of photon-stimulated to thermally-stimulated electrons will be very low. If the cathode is made smaller, the number of thermionically emitted electrons will decrease while - up to a point - the number of photo-emitted electrons (for the same light flux) will be the same. Thus in the case of photoemissive surfaces at least, it would seem that the significance of thermal noise effects - from the point of view of detecting a given light flux - would be proportional to the size of the cathode. Although the mechanism of photodetection in the retina is quite different, it seems plausible that similar considerations apply, i. e., that the signal to noise ratio is related inversely to the size of the photodetectors. It may be noted that an image orthicon TV camera is a fairly close analog of the eye - in terms of size, number of grain elements, and sensitivity. In this camera tube the basic photodetectors can be considered as the very small elements of the photosensitive screen corresponding to the grain size of the TV picture.

It seems that the sensitivity of the eye, although probably superior to most inanimate detectors, is not unique and can be approached and duplicated if necessary, although in equipment larger and generally less convenient than that of the eye. As suggested above, the duplication in inanimate equipment of the extremely small size of the rods and cones and their interconnections may represent an approach to fully duplicating all the good features of the eye - e. g., high sensitivity at room temperatures, small size, low power consumption, etc.

There are several other aspects of the visual photodetecting system that may be considered for future exploitation.

1. The photo-chemical-electrical process associated with the rods and cones:
 - a. with the actual chemicals rhodopsin and iodopsin
 - b. with other complex chemicals.
2. Mass connections to the individual photoreceptors in a pattern detector, by electronic wires or "neuristor" wires. This is in contrast to the time division scan over all the elements by a single beam, as in the image orthicon TV tube.
3. Data processing of the detail in an image before transmission.



The type of benefits to be looked for as a result of this kind of exploitation are the development of visual light and pattern detectors that:

1. **are light, compact, very sensitive and require little power to operate.**
2. **are adaptable over a wide range of illumination.**
3. **are capable of extracting useful information about the pattern being observed.**

This is particularly relevant to the very pressing need for automatic character and script recognition devices. The absence of such devices is a serious bottleneck in the utilization of even existing electronic data processing equipment.

4.3.2 Oculometry

4.3.2.1 The Techniques of Oculometry

A considerable amount of work has been done in recent years on the development of techniques for the measurement of the orientation of the axis of regard (Ref. 9). All the work described in the literature, so far discovered, pertains to devices constructed for the purposes of physiological and psychological research. The equipment produced for these purposes has, in the main, only been suitable for laboratory experiments.

The following are the principal techniques of oculometry:

1. **Electro-oculometry (Refs. 18, 19).** A small potential difference exists between the cornea and retina of the eye. The effect of this potential difference spreads out into the surrounding tissue and can be detected by electrodes placed on the skin near the eyes. As the eyeball rotates, the orientation of the dipole axis changes and the voltage measured by the electrodes changes correspondingly. The sensitivity is of the order of $20 \mu\text{v}/^\circ$. The accuracy of the system, limited principally by spurious skin potentials, is said to be about 1° .

2. **Corneal Reflection.** The cornea is a nearly spherical surface of radius about 8 mm. The center of rotation of the eye is about 13 mm. from the surface of the cornea (Fig. 4.7). Thus rotation of the eye will cause an image reflected in the cornea to move. However lateral motion of the eye - for example due to lateral motion of the head - will also cause the image reflected in the cornea to



move. In fact, a lateral displacement of only 0.01 mm. is said to produce an effect equal to a pure rotation of 8 minutes of arc. Thus even with the most rigid clamping of the head, the system cannot be considered very accurate.

3. Contact Lens Reflection (Ref. 20). By fixing a plane mirror on a contact lens and projecting a collimated beam of light on to it, image motion is obtained only by rotation of the mirror - and hence the eye. Lateral motion of a plane mirror does not deflect the image of a collimated beam. This method is extremely sensitive, a displacement of one second of arc can be detected. Slippage of the lens over the eye is said to occur during large motions of the eye.

4. Photoelectric (Ref. 21). In this method an image of the edge of the iris and part of the cornea is projected onto a slit behind which is a photocell. As the eye moves, the relative amounts of the images of the cornea (light) and iris (dark) falling on the photocell changes. Thus the output of the photocell is proportional to eyeball rotation. The whole apparatus has to be clamped to the head. The accuracy of the system is said to be about 15 minutes of arc.

5. Retinal Observation (Ref. 22). All other methods of oculometry measure eyeball rotation. This method directly measures the displacement of the retinal image relative to the retina, which is usually the actual measurement that is required. Although the motion of the retinal image is largely determined by eyeball rotation, the possibility of other intra-ocular effects on the position of the image on the retina cannot be ignored if great accuracy is to be obtained.

In the retinal observation method an external image is formed of the blind spot of the retina where the optic nerve enters the eye. Blood vessels stand out clearly against the background and are used to activate a photoelectric device. As the image of an external scene on the retina moves over the retina, the image of the retina will move relative to the external scene. This motion is detected by the photocell. The accuracy of the equipment described in reference 20 is stated as being of the order of 10 seconds of arc, over a range of about 4° of eye motion.

4.3.2.2 Applications of Oculometry

Human visual acuity is concentrated sharply around the axis of regard and consequently the oculomotor apparatus of the eye - a precise servo control system - is a very important part of the visual sense; it is particularly important in the perception of spatial relations.



These facts suggest that,

1. spatial information derived by the visual sense can be extracted from a human, or animal, by the technique of oculometry.
2. a TV system may be constructed in which the grain size matches the visual acuity of the eye over its whole angular range. The result would be a practical system of HI-FI TV.

These general possibilities and others are developed in the succeeding parts of this section.

The systems described for the various applications involve interconnections of an oculometer (ARD - axis of regard detector); a device to control the orientation of the optical axis (VOA - variable optical axis); and a zoom lens (VM - variable magnification).

Of these three devices only the ARD is discussed in detail in the later sections as the other two involve no new principles.

The VOA unit will be required to operate over a field of about 90° corresponding to the possible range of motion of the eye. It must be capable of following the fastest motion of the eye to within 1° (assuming a 1° dynamic range of the oculometer). Thus it must be able to attain a rate of rotation of $500^\circ/\text{sec.}$ in $1/500 \text{ sec.}$ This corresponds to an angular acceleration of $250,000^\circ/\text{sec.}^2$ or about $4,000 \text{ radians/sec.}^2$. Assuming that the VOA unit is simply a rotatable mirror of 1 oz. mass and a radius of gyration of $1/4 \text{ inch}$ the torque needed to attain this angular acceleration is,

$$\begin{aligned}
 &= \frac{1}{32.2} \frac{1}{16} \frac{1}{144} \frac{1}{4^2} 4 \times 10^3 \text{ ft. lbs.} \\
 &= \frac{1}{32.2} \frac{1}{12} \frac{1}{4^2} 4 \times 10^3 \text{ in ozs.} \\
 &\approx \frac{1}{10^2} \frac{1}{16} 10^3 \approx 0.6 \text{ in ozs.}
 \end{aligned}$$

which is evidently quite feasible.



4.3.2.2.1 Command Control Systems

Some weapon systems, and their guidance units, are automatic while in others a human is "built in" as an integral part of the system. For example, once it is fired, the Sidewinder missile is automatic, whereas the SS-10 anti-tank missile must be controlled by a human operator throughout its flight. It is instructive to consider the reasons for including a human in any weapon system.

A human operator is used to control an anti-tank missile because a reliable **method of automatically identifying tanks, against the normal battlefield background,** has not been found. Compared to a tank on a battlefield, an airplane in an empty sky stands out "like a sore thumb", and it has proved relatively easy to develop an automatic means of identifying a plane and thus mechanizing an automatic missile guidance system.

Although it seems difficult for an automatic device to recognize a tank as compared to a plane, a human does it relatively easily. The reason, of course, is that the human brain performs the detection function by using its phenomenal power of character and pattern recognition, and not by a simple "yes-no" quality, as does an IR or microwave guidance system.

The magnitude of the character or pattern recognition power of the human brain can be appreciated by attempting to formulate a computer system to simulate even one aspect of it. For example, we are capable of recognizing, in a fraction of a second, any one of possibly thousands of our friends and acquaintances should they pass by in a busy city street. Presumably the brain must be constantly attempting to correlate the observed visual patterns with the contents of the visual memory. Such a capability is clearly well beyond the present state of computer technology.

It is evident that the use of a human in the SS-10 weapon system is in fact an example of the use of a bio-sensor - in this case the human sense of pattern or form. However, in the SS-10, and other similar command control systems, the human operator is required to do more than recognize the target. He must observe the missile, see its position relative to the target, judge - on the basis of past experience - how much to move the control stick, and finally actually to move the stick. In more refined systems the operator is required only to manually track



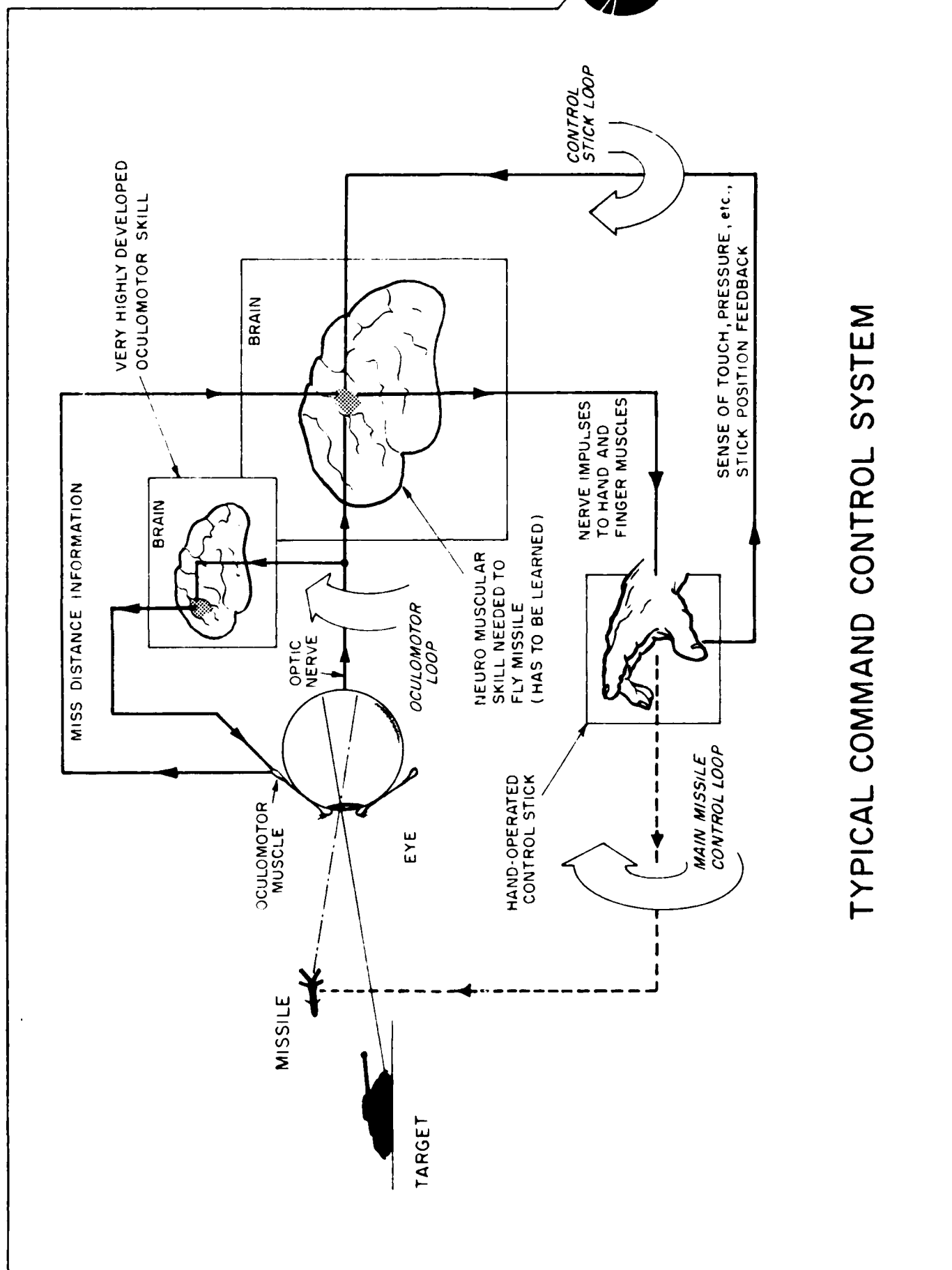
the target. However, in either case, he is called upon to do more than is necessary, i. e., it is absolutely necessary that he recognize the target - simply because no other device exists that can do it - but it is not necessary to employ a human to perform the other functions particularly that of moving a control stick or any other manual activity.

As will be shown, the technique of oculometry will allow the human operator in a command control system either to confine his activity to locating the target to be attacked, or alternately to locating the target and the missile. In neither case will he be required to perform any manual control functions.

A typical line of sight command control system is shown in Figure 4-12. It will be seen that three feedback loops are involved, the oculomotor loop, the main missile control loop, and the control stick loop. The oculomotor loop controls the orientation of the axis of regard; it is a precise control system that plays a very important part in the total visual sense. Spatial information comes into the consciousness, not so much from the optic nerve, but from the oculomotor muscles and nerves. The oculomotor loop has been shown in Figure 4-12 containing a section of the brain responsible for the very highly developed oculomotor skill. (As an indication of the degree of skill involved, it may be noted that the eyes can be moved at up to $500^{\circ}/\text{sec.}$, over a range of about 45° , with an accuracy of better than 10 minutes of arc).

The spatial information entering the consciousness is the basis on which control stick motion is derived. It is to be noted that the generation and execution of control stick commands by the consciousness is a skill that has to be learned. In the first place, an optimum functional relationship exists between observed missile deviation and control stick deflection, and in general, this is not a simple time invariant scalar function, so this has to be learned from experience. Secondly, the movement of a control stick is not an every day muscular activity - as is motion of the eyeball - and this too has to be learned - just as a typist or pianist had to learn a particular muscular skill.

The control loop responsible for the motion of the control stick has only a very poor position feedback path via the senses of touch and pressure, etc. and thus the operator has a poor estimate of how much he has actually moved the stick at any time. He must, therefore, inevitably generate noise as well as effective



TYPICAL COMMAND CONTROL SYSTEM



control signals. In the oculomotor loop, on the other hand, an extremely good feedback path exists through the optic nerve. It is well known that the performance of a position servo is directly related to the quality of the position feedback path.

Thus, in the command control system shown in Figure 4-12, the oculomotor loop that generates spatial information to be sent to the consciousness, is an efficient and essential part of the system. However, the generation from this spatial information of control stick commands, and the execution of these commands, is unsatisfactory in that:

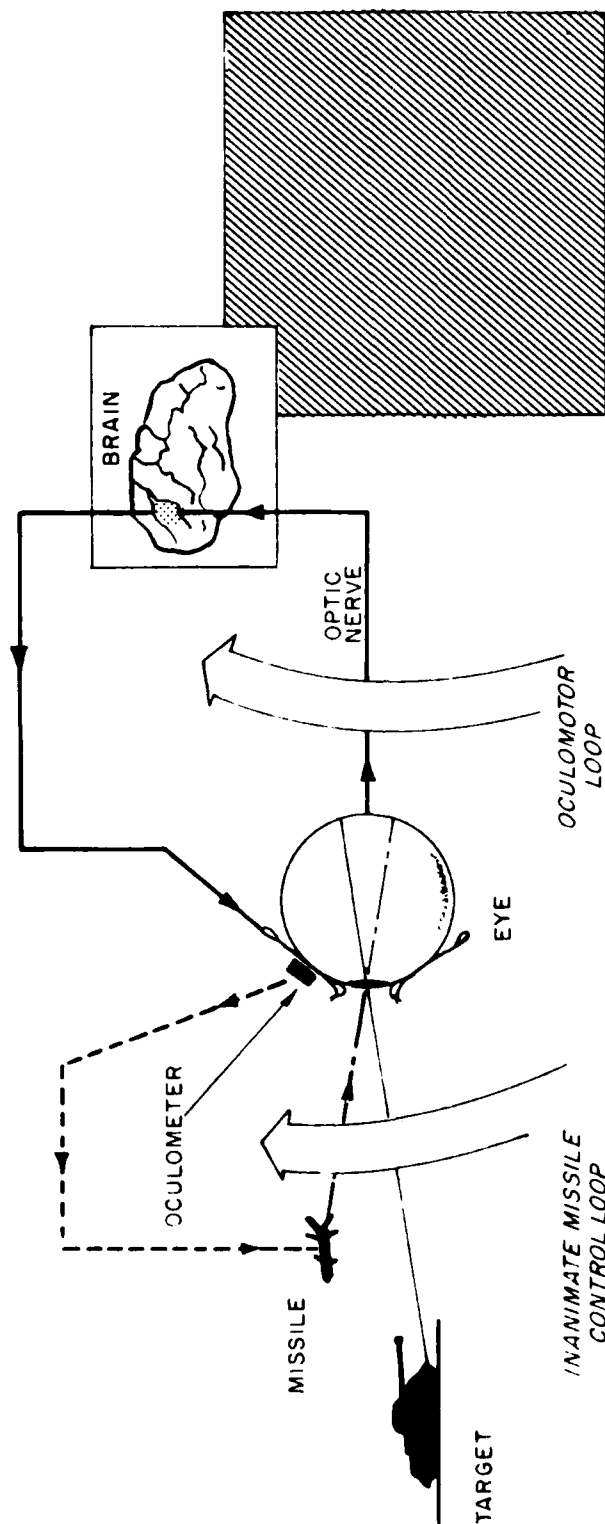
1. A skill has to be learned.
2. The control stick loop is noisy, due partly to a poor feedback path.
3. Even when the control skill has been learned, a transfer function of very limited bandwidth and response time is interposed between the spatial information coming from the oculomotor muscles and the electrical output of the control stick.

The technique of oculometry will permit the extraction, from the operator, of the missile/target spatial information at the same time, and in the same form as it is supplied to the consciousness (Fig. 4-13). In this way, the undesirable features of a conventional human operator command control system can be avoided.

There is a limitation involved in the extraction of spatial information by the technique of oculometry. Within an angular width of about 10 minutes of arc from the center of the fovea, the eye motion is random and attention can be fixed on any point within this area without moving the eyeball. Thus, when the missile and target are within 10 minutes of arc of each other, the spatial information is derived essentially from the optic nerve and not, as shown in Figure 4-12, from the oculomotor loop. This, of course, does not alter the fact that the human operator is being used inefficiently in the set-up shown in Figure 4-12. The limitations, discussed earlier, involved in the command and execution of control stick motion still apply. However, the arrangement shown in Figure 4-13 will not work at all when the missile and target are so close. When it is recalled that visual acuity is only of the order of 1 minute of arc, it can be appreciated that the breakdown of the system shown in Figure 4-13 occurs quite near to the threshold of acuity.



COMMAND CONTROL SYSTEM WITH OCULOMETER





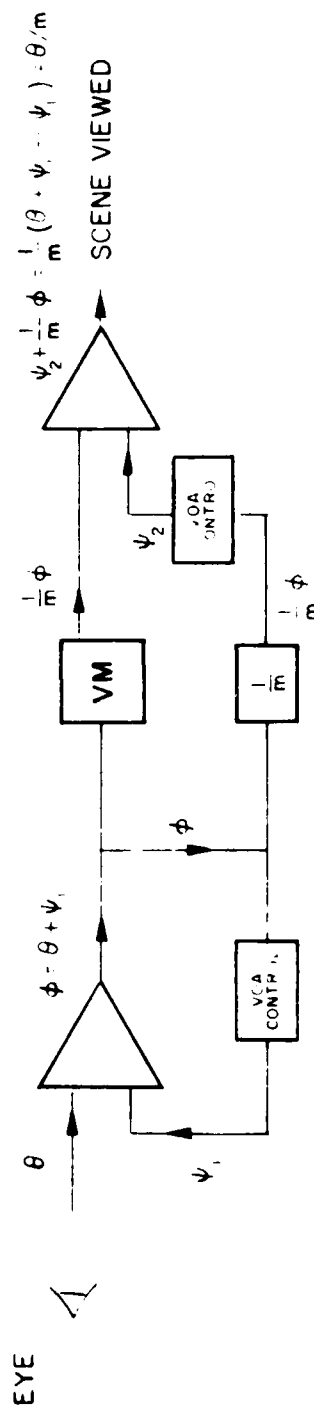
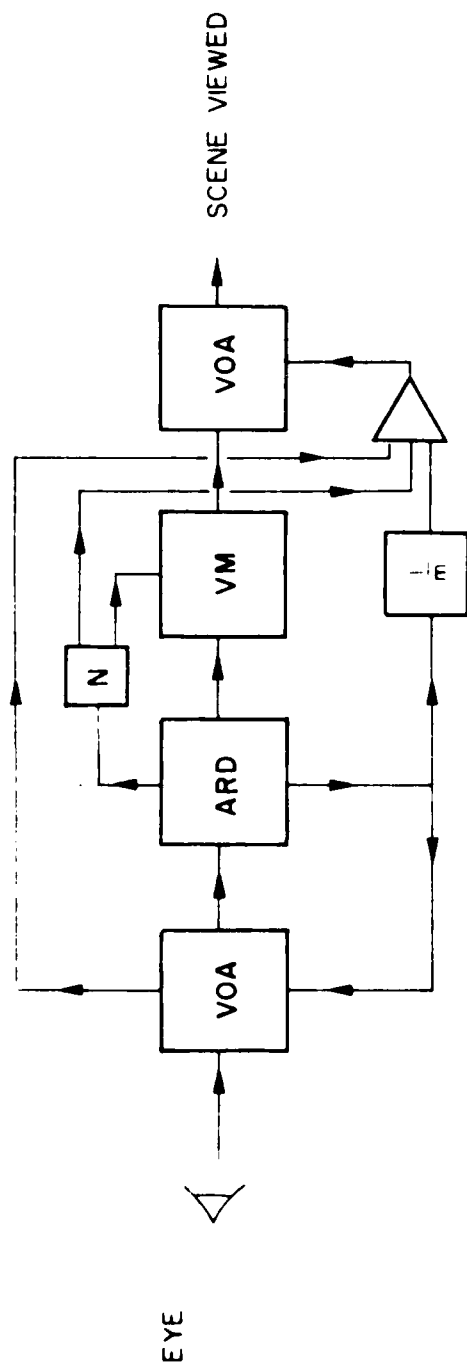
The solution to the difficulty lies in the use of a continuously variable magnification device - i.e. a zoom lens - shown as VM in Figure 4-14. The ARD unit acts as a null sensor in a loop associated with the VCA unit at the eye, and the two units together comprise a wide range oculomotor. The eye motions are analyzed by a network, N; if the short term variable saccadic motion is less than say, 10° , the magnification is smoothly increased, if greater than 10° the magnification is decreased. The mean value of the eye motion may also be sensed and used to reorient the VOA unit. In this way the detail of interest in the viewed scene is kept centered in the field of view and its angular size is kept up to 10° by the action of the variable magnification unit (see Fig. 4-15). The missile/target displacement vector is taken from the oculomotor system and processed through an optimum control network for transmission to the missile as guidance information.

With a 10° "signal" and a "noise" of 10 minutes (i.e. the magnitude of the random, involuntary motion of the eye) the signal to noise ratio is 60:1.

An upper limit must, of course, be set to the magnification, since otherwise it would tend to infinity when the missile was on the line of sight. For a missile attacking a tank, the required impact accuracy is about 2 ft. Taking 20 minutes of arc as a conservative estimate of the oculomotor noise level, and specifying that for high kill probability at a maximum range of two miles 1 ft. should correspond to 20 minutes of arc, the maximum magnification required is seen to be 60. The operator's view at impact, at 2 miles range, with a magnification of 60, is shown in Figure 4-16.

Since the optical system is self-aligning, it may be hand held or clamped on a tripod. The system can be designed to operate in the presence of vibration that would normally be intolerable with a magnification of 60. Given the amplitude/frequency spectrum of vibration, the performance of the servo elements in the arrangement shown in Figure 4-14 can, within limits, be specified to attenuate the vibration to an acceptable level.

In the system described above, the operator is required to perceive the spatial relation between missile and target. In order to do this, he will be forced to flick his eyes from one to the other, since the variable magnification device



$$\psi_1 = \frac{C\phi}{ap^2 + bp}$$

m = MAGNIFICATION OF VM UNIT

$$\psi_2 = \frac{-C(\frac{1}{m}\phi)}{ap^2 + bp} = -\frac{\psi_1}{m}$$

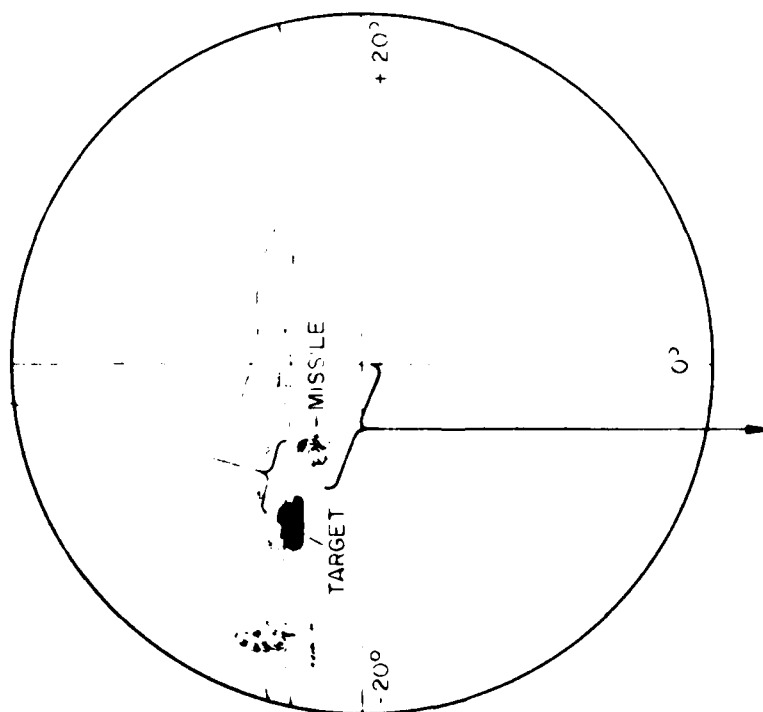
BLOCK SCHEMATIC AND FLOW GRAPH OF OCULOMETER COMMAND CONTROL SYSTEM

ALLIED RESEARCH ASSOCIATES, INC.

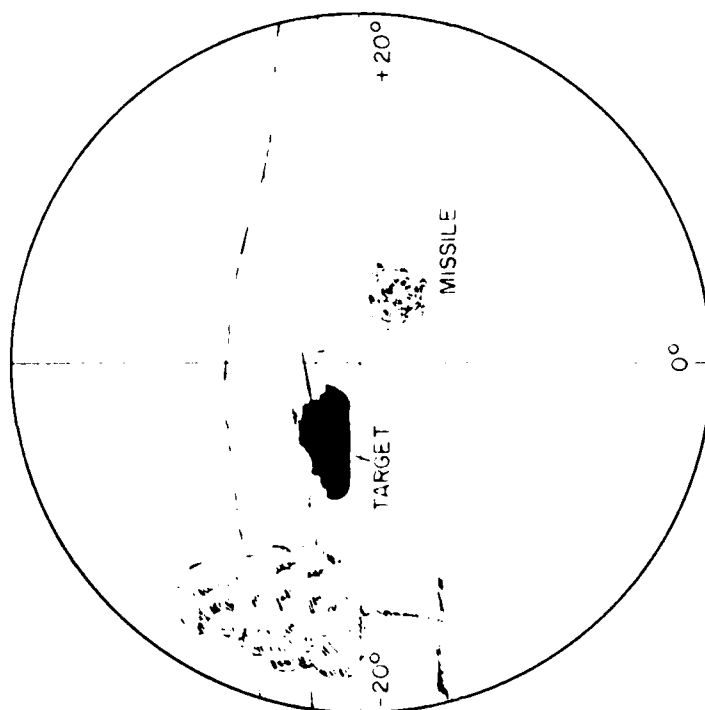


BOSTON, MASSACHUSETTS

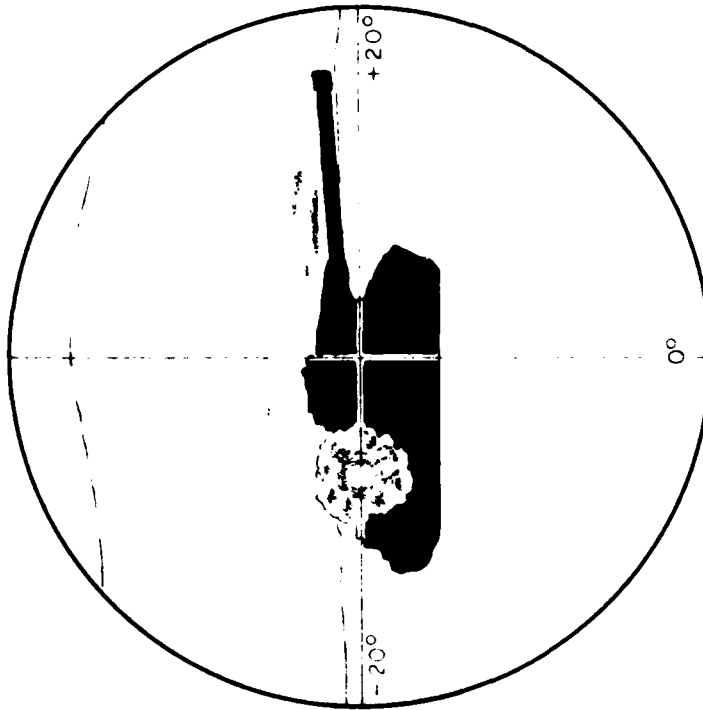
MAGNIFICATION INCREASED UNTIL THIS
DISTANCE SUBTENDS 10° AT EYE



OPTICAL AXIS REORIENTED UNTIL THIS
DISTANCE SUBTENDS ZERO ANGLE AT
EYE



OCULOMETER COMMAND CONTROL SYSTEM: OPERATORS VIEW BEFORE AND AFTER CORRECTION



OPERATORS MAGNIFIED VIEW OF TARGET AND MISSILE
AT IMPACT, RANGE 2 MILES



will ensure that missile and target are separated widely enough to prevent spatial perception by the retina. The missile/target displacement vector is obtained by detecting, electronically, this saccadic motion.

In this system, the operator has not only to recognize the target, but also the missile. The latter function can be done automatically, but it may be advantageous to use the operator to recognize the missile in order to reduce equipment complexity with little sacrifice of system performance.

In an alternative application of oculometry, the position of the missile may be sensed automatically by IR radiation. The operator's task may then be reduced to the essential minimum of simply observing the target. In order to avoid bore sighting problems, the IR and visual optics should be arranged so that they share a common optical axis. The principle of operation will be very similar to the system shown in Figure 4-14. The orientation of the viewing axis will be controlled electronically so that the visual and IR images are symmetrically placed in the center of the field. The magnification will be adjusted in the same way as shown in Figure 4-14.

4.3.2.2.2 Television

Although a domestic TV picture has quite good definition - i.e. when viewed correctly, it does not reveal its "grain" structure - the total visual impression gained by watching TV is entirely different from the impression gained by watching the televised scene directly. The difference stems mainly from the very small angle of view that is possible with domestic TV.

The domestic TV picture can be considered as a mosaic of 500 rows each containing 500 dots or grain elements. Normal visual acuity is of the order of 1 minute of arc so that when viewed so as not to show its grain structure, a TV picture cannot have an angular width of more than 8° . With a screen size of 15" X 15" this corresponds to a viewing distance of 9' i.e. a normal viewing distance.

The normal field of view of one eye is 120° - 150° and with both eyes it is in excess of 180° . Thus, in terms of solid angle, a TV picture is only 0.2% of the size of the total visual field. In domestic TV the limitation imposed by the very small angular width can be overcome to some extent by efficient studio production



techniques, such as switching the transmission from one camera to another. This is analogous to the way the oculomotor apparatus of the eye normally switches the gaze from point to point in a visual scene. However, even under the best conditions, the sensual impression gained by the viewer is a poor substitute for the impression derived by being actually present at the televised scene. In many non-entertainment applications of TV it is not practical to have a number of cameras and a producer at the televised scene, so that the narrow field of view is then a very great limitation from the point of view of the total visual impression that can be gained.

This present application of oculometry is concerned with an entirely different concept of TV. This is to create a TV system which would completely duplicate the visual impression gained by an observer actually present at the sight of the TV camera. The aim, in effect, will be to engineer a remote extension of the visual sense organ by devising an efficient match between the eye and the TV system.

As discussed earlier, our contact as human individuals with the external environment is almost totally dependent on our sense of vision and, to lesser extent, of sound. Thus, the existence of a HI-FI TV system of the type being considered would, in effect, enable us to experience an environment, without actually being present. It would be as if our optic and auditory nerves were hundreds or thousands of miles long instead of only a few inches. What we customarily call being "present" in an environment really means that our sense organs, i.e. eyes, ears, etc., are present. When electronic extensions of the two most important sense organs are available it will not be necessary for our bodies to be physically located at a place in order to experience the sensations of being "present" at that place.

Numerous applications can be envisaged for HI-FI TV:

1. Exploration of moon and planets.

There are many difficulties associated with manned space exploratory missions. However, unmanned missions with either automatic or remotely controlled equipment will, in general, be very much less useful. The advantages of having a human astronaut on the planet being explored, rather than just inanimate equipment, are that a human can perform a very wide range of mechanical tasks -



putting soil under a microscope, operating scientific instruments, etc. - and that he is intelligent and adaptable. Automatic equipment - e.g. a soil analyzer - may fail if the conditions on the planet are not like those expected by the designer of the equipment. However, a human astronaut, equipped with only a few simple tools, would be able to react effectively to a wide range of unexpected situations.

If it were not for the transmission delay, caused by the finite velocity of electromagnetic waves, HI-FI TV could obviously be usefully applied to earth based planetary exploration. However, even to the moon, the delay involved in the round trip communication link is $2\frac{1}{2}$ secs. and is certainly too great to allow for the successful operation of a truly HI-FI TV system. However, the HI-FI TV system could be employed on manned missions to orbit the various planets. Instead of landing the non-expendable astronauts, a HI-FI TV system could be sent down. This would allow the orbiting astronauts to experience the same visual sensation as actually being on the planet, without incurring the enormous weight penalty of the return booster. This scheme would allow for lunar and planetary exploration to be undertaken with boosters that will soon be available. The much larger boosters needed for a manned landing on the moon or planets will probably not be available for a number of years. Conventional TV is being considered for earth based unmanned robot exploration of the moon in the immediate future (Ref. 4). The round trip communication delay of $2\frac{1}{2}$ sec. is a limitation that has to be accepted. In the system described in Reference 4, the TV camera has a field of view of 45° , the frame repetition rate is 1 per second, the video bandwidth 200 kc, and the transmitter power 3.16 watts. It is estimated from these figures that the grain size in this TV system is 6 minutes of arc. The oculometrically controlled TV system, described later in this section, would require a bandwidth of only 3 kc for 1 minute of arc grain size at the center of the field, and a picture repetition rate of one frame every $2\frac{1}{2}$ seconds. Thus, although it cannot be considered in this application as a duplicate of the visual sense - because of the $2\frac{1}{2}$ second communication delay - oculometric TV can nevertheless be considered for its economy of bandwidth.

The use of HI-FI TV alone would permit a form of passive lunar or planetary exploration. However, by incorporating such a TV system into an advanced form of robot, an almost fully realistic form of active manned exploration could be



achieved. Such a robot would be equipped with analogs of the human senses (e.g. vision) and motor output (e.g. limb motion). It would be controlled by a human whose senses would be completely coupled, via communication channels, to those of the robot - so that the human would receive virtually no sensory input from his local environment. The motor activity of the human controller would be monitored and transmitted to the robot, and this information used to command corresponding motion of the robot. Thus the human controller would be in contact, not with his own environment, but with that of the robot. It is emphasized that the fidelity of the robot sensory detectors, man-machine coupling, etc. could be such as to provide the robot controller with a second "body" which would be under the control of the controller's consciousness in almost the same way as is his normal body. The robot would thus appear to have human or near human dexterity and skill - together with the intelligence and initiative of the human consciousness controlling it. The effect of the round trip communication delay of 2 - 1/2 seconds could be mitigated by including certain closed loop control circuits in the robot - e.g. for postural stability.

The robot could be used for lunar and planetary exploration. In this way many of the advantages of manned space missions could be combined with the advantages of unmanned missions. In particular the problems of biological contamination, either of the earth or the explored planet, are avoided.

2. Meteorological and strategic surveillance from earth satellites by means of a TV monitor on the earth. The following advantages are postulated.

- a. Many people, military, executive, and scientific, while on the earth could experience a visual sensation equal to, or better than, that experienced by orbiting astronauts.
- b. No life support problem.
- c. The need to ferry supplies and personnel to and from orbiting space stations would be largely eliminated. Maintenance could be carried out, in most cases, by remote control.
- d. The Van Allen radiation belts would impose less restrictions on the satellite orbit.
- e. The acknowledged search capability of the human eye could be employed in a missile launching detection system. Observers could be changed frequently - as is the practice for radar observers.

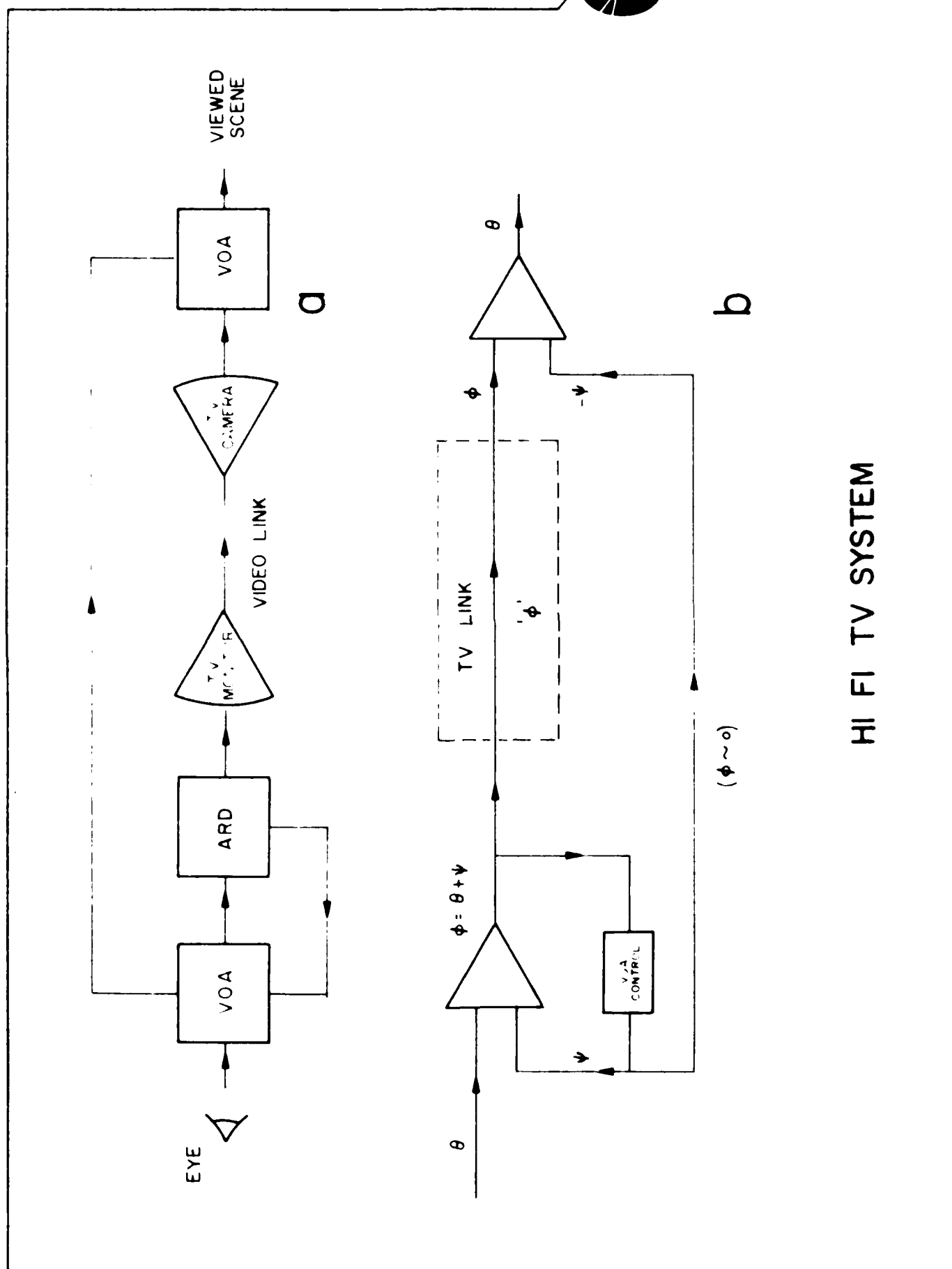


- f. In conjunction with an advanced robot system - the satellite (and other vehicle borne robots) could be maintained and repaired.
3. Navigational aid for aircraft pilots. With an IR sensitive camera, bad weather visibility could be improved.
4. Battlefield surveillance for army commanders, artillery support, etc. With separated cameras, binocular vision could be provided with a greatly enhanced sense of depth. Alignment problems normally associated with such binocular devices would be eliminated.
5. Miscellaneous weapon systems in which a robot viewer could be considered more expendable than military personnel.
6. Guidance techniques involving missile borne TV would become more attractive with the very much lower bandwidth needed with oculometrically controlled TV.
7. When satellite communication systems become operational, the availability and cost of long distance wideband communication circuits may be such as to permit the use of HI-FI TV for such purposes as business conferences; instead of all the participants travelling to a common location, they could use a HI-FI TV network to derive an equivalent sensual impression.

The general form of a HI-FI TV system is shown, in Figure 4-17a, as an interconnection of an ARD and two VOA units, a conventional TV system and a communications link. A flow graph of the angle between the axis of regard and the central axis is shown in Figure 4-17b. The eye is shown with its axis of regard displaced by an angle θ from the central axis, and the VOA units deflecting the optical axis by an angle ψ , where $\psi \sim -\theta$. At the TV link, the axis of regard is at angle $\phi = \theta + \psi$ to the central axis, where $\phi \sim \text{zero}$.

The TV monitor projects into the eye via a VOA unit. This latter is driven by a signal from the ARD unit to a deflection designed to bring the axis of regard onto the central TV axis, no matter in what direction the eye is actually looking. Thus let ψ be the deflection of the VOA unit so that the axis of regard is at ϕ to the central axis throughout the TV link. ϕ will be measured by the ARD unit: let ψ be controlled according to the law:

$$(a p^2 + b p) \psi = -c \phi$$



HI FI TV SYSTEM



(i. e. the torque applied to rotate the VOA axis will be proportional to ϕ and a rate signal $\dot{\phi}$)

$$\therefore (a p^2 + b p + c) \dot{\phi} = -c \theta$$

i. e. $\dot{\phi} \rightarrow -\theta$ as $t \rightarrow \infty$

and

$$\phi = \theta + \psi = \left(1 - \frac{c}{a p^2 + b p + c}\right) \theta$$

i. e. $\phi \rightarrow 0$ as $t \rightarrow \infty$

At the remote end, the TV camera is pointed to the scene being viewed via another VOA unit which is slaved to the VOA unit in the receiver terminal. Thus, during transmission through the TV link the axis of regard is kept to an angle ψ from the central axis, where $\psi \rightarrow 0$, but at the transmit terminal it is restored to θ , the value at the eye.

In order to make economical use of bandwidth, the resolution of the TV picture must be made a function of the angular displacement from the central axis. This could be done either electronically or optically.

It is possible to discuss the feasibility of the HI-FI TV system in terms of the following general aspects of the system.

1. The bandwidth required in the TV communications link.
2. The static accuracy and linearity of the VOA and ARD units.
3. The dynamic response characteristics of the VOA and ARD units.
4. Mechanical and optical details such as the coupling of the eye to a wide angle optical system.

The bandwidth requirements of the link can be computed from the number of grain elements and the picture repetition rate. If a grain structure equal to a normal TV picture viewed so that the lines are just resolvable is to be obtained, then 10,000 grain elements are required. There is a possibility that a picture sequencing repetition rate as low as 10 per second could be employed by synchronizing the picture sequencing with the occurrence of saccadic motion of the eye. This system would require a bandwidth of 70 kc. A system employing ten times as many grain elements - which would correspond to the pupil diffraction limit on visual acuity - and a picture repetition rate of 30 per sec. would require a bandwidth of 2 mc. Thus, it is clear that bandwidth requirements - even for a very quality system - are modest.



The two VOA units must have a static accuracy, or linearity, sufficiently good to allow them to follow each other to within about 1° . A larger error than this would result in noticeable distortion in the visual field, as the eye moved from place to place.

The ARD and VOA units must have a dynamic response characteristic adequate to insure:

1. that the ARD unit is not forced out of its operating range.
2. that 1/10 second or more after a saccade has occurred, the axis of regard does not deviate more than about 10 minutes of arc from the central, high definition, axis of the TV system.

Assuming that the fastest motion of the eye is $500^\circ/\text{sec.}$, the maximum deviation ϕ that will occur is

$$\phi = \frac{500}{p^2} \left(1 - \frac{c}{a p^2 + b p + c} \right)$$

If ω_n is the natural frequency of the second order system $(a p^2 + b p + c)$ and if this is properly damped, it can be shown that the maximum value of ϕ is of the order of $500/\omega_n$. If this is to be 1° , then $1/\omega_n$, the response time of the system, must be about 1/500 sec. When the ramp motion $500/p^2$ stops, the maximum value of ϕ will be about 1° . With a response time of 1/500 sec. this value of ϕ will fall to 10 minutes of arc in considerably less time than the 1/10 sec. specified.

It has been shown that a HI-FI TV system would require a relatively low bandwidth transmission circuit, somewhere between 70 kc and 2 mc. The ARD and VOA units should have a response time of the order of 1/500 sec., and the ARD unit a dynamic range of $\pm 1^\circ$. The linearity of the VOA units should be adequate to ensure that they follow each other to within 1° .

It is concluded that the HI-FI TV system is feasible: development work would be required particularly on

- a. the variable resolution scanning system
- b. the coupling of the eye to a wide angle optical system
- c. the possibility of synchronizing the picture sequencing to the saccadic motion of the eye, in order to reduce the picture repetition rate, and thus the required bandwidth.



4.3.2.2.3 Search Light Illumination

In this application, the axis of an illuminating light beam is kept parallel to the axis of regard of the person using the illumination to view a dark scene. In this way the available light energy can always be kept concentrated so that most of it falls on that part of the scene where it is most needed and is most effective - i.e. at the point which forms an image directly over the fovea of the observer.

It has been seen that foveal acuity is more or less governed by the number of light quanta available. However, under normal illumination peripheral acuity is much poorer than the quantum limit. No quantitative information has been found concerning the relationship between peripheral acuity and illumination. Peripheral vision is, in the dark adapted eye, very sensitive. The acuity does not seem - objectively speaking - to be very much less than in normal daylight. Thus it appears possible that a considerable reduction could be achieved in the power needed to (volume) illuminate a given scene by arranging for that part of the scene observed by foveal vision to be illuminated more strongly than that observed by peripheral vision. In other words the quantum flux available would be matched to the resolving power of the eye over the whole field of view. There are many factors that could limit the extent to which it would be practicable to shape the beam intensity - e.g. finite response time of the search light axis servo, relationship between foveal light intensity and peripheral dark adaptation, internal light reflection in the eye, etc. On the basis of the acuity curve given in Figure 6 it appears that, theoretically, a very considerable power reduction could be achieved. Along the axis of regard, the resolution is one minute of arc, (under normal illumination). As has been seen, this one minute of arc in the scene contributes a quantum flux of the order of 1 quantum per summation time at the eye. Under the same illumination, the resolution 10^0 off the axis of regard is 100 times less. Thus the area of resolvable detail is 10^4 times greater and will contribute a flux of the order of 10^4 quanta per second at the eye. The quantum inefficiency (at normal illumination) of peripheral vision is now apparent. What is suggested here is that the intensity of illumination of these peripheral areas be reduced, so that - with dark adaptation - the quantum efficiency may be improved to that of the fovea. A calculation, similar to that given in Section 4.1.3 on the information content of the retinal image, shows that a theoretical maximum power saving of the order of 10^4 times would be possible.



4.3.2.2.4 Stabilization of Optical Instruments

It is well known that it is difficult to use hand held binoculars that have a magnification much greater than about 10 times. The natural tremor in the hand is sufficient to cause the image to move so rapidly that the eye cannot follow it. This difficulty could be partly overcome by arranging for the fine pointing of the optical axis to be servo controlled by a built in oculometer. This arrangement would not only filter out the hand tremor, but also assist the eye ball in keeping up with any residual tremor in the image.

4.3.2.2.5 Miscellaneous Applications

1. It has been shown (Ref. 3) that when the effect of the normal involuntary eye motion is doubled, acuity is improved and steady fixation becomes effortless. In effect, the gain of the normal feedback loop controlling fixation is increased. By the use of an ARD and VOA unit other gains and phase shifts could be introduced.
An optical viewing device incorporating this principle might be of value in manual-optical search and tracking tasks.
2. With a suitable arrangement of VOA and ARD units the axis of regard of one observer could be slaved, to within about 10 minutes of arc, to that of another. The passive observer would be able, in fact compelled, to see what the active observer was looking at. It would be impossible for him to allow his gaze to wander. Applications can be foreseen in education, indoctrination, and crime detection.
3. Physiological and psychological research.
4. With the development of suitable oculometric devices, the visual apparatus of animals could be controlled. The visual sense is the principal available contact with the brain, so that oculometry could provide the basis for more efficient linking of animal brains to machines.

REFERENCES - SECTION 4

1. Best, C. H., Taylor, N. B. Physiological Basis of Medical Practice. 5th Edition, Williams and Wilkins Co., Baltimore, (1950).
2. Pritchard, R. M. Stabilized Images on the Retina, Scientific American June 1961.
3. Riggs, L. A., Ratliff, F., Cornsweet, J. C., Cornsweet, T. N., The Disappearance of Steadily Fixated Visual Test Objects. Journal of the Optical Society of America, Vol. 43, p. 495, June 1953.
4. Roving Lunar Vehicles Pt. II. Aviation Week. p. 71, October 9, 1961
5. Pirenne, M. H., Denton, E. J. Accuracy and Sensitivity of the Human Eye. Nature, Vol. 170, p. 1039, 1952.
6. Schroeder, A. C. Theory on the Receptor Mechanism in Color Vision. Journal of the Optical Society of America. Vol. 50, p. 945, October 1960.
7. Keesey, U. T. Effects of Involuntary Eye Movements on Visual Acuity. Journal of the Optical Society of America. Vol. 50. p. 769, August 1960.
8. O'Brien, B. Vision and Resolution in the Central Retina. Journal of the Optical Society of America. Vol. 41, p. 882, December 1951.
9. Shackel, B. Review of the Past and Present in Oculography. Medical Electronics, Proceedings of the II International Conference on Medical Electronics at Paris, 1959.
10. Hebbard, F. W., Marg, E. Physiological Nystagmus in the Cat. Journal of the Optical Society of America. Vol. 50, p. 151, February 1960.
11. Krauskopf, J., Cornsweet, T. N., Riggs, L. A. Analysis of Eye Movements During Monocular and Binocular Fixation. Journal of the Optical Society of America. Vol. 50, p. 572, June 1960.
12. Riggs, L. A., Armington, J. C. Ratliff, F. Motions of the Retinal Image During Fixation. Journal of the Optical Society of America. Vol. 44, p. 315, April 1954.



13. Ford, A., White, C. T., Lichtenstein, M. Analysis of Eye Movements During Free Search. Journal of the Optical Society of America. Vol. 49, p. 287, March 1959.
14. Cornsweet, T. N. Determination of the Stimuli for Involuntary Drifts and Saccadic Eye Movements. Journal of the Optical Society of America. Vol. 46, No. 11, p. 987.
15. Shortess, G. K., Krauskopf, J. Role of Involuntary Eye Movements in Stereoscopic Acuity. Journal of the Optical Society of America. Vol. 51, p. 555, May 1961.
16. Ludvigh, E., Miller, J. W., Study of Visual Acuity During the Ocular Pursuit of Moving Test Objects: I and II. Journal of the Optical Society of America. Vol. 48, p. 799, November 1958.
17. Maturana, H. R., Lettvin, J. Y., McCulloch, W. S., Pitts, W. H. Anatomy and Physiology of Vision in the Frog. Journal of General Physiology, July 1960. Vol. 43, Number 6 Supplement, p.129.
18. Shackel, B. Note on Mobile Eye Viewpoint Recording. Journal of the Optical Society of America. Vol. 50, p. 76 B. August 1960.
19. Shackel, B., Sloan, R. C., Wass, H. J. J., Electronics. Vol. 31, p. 36, January 31, 1958.
20. Ratliff, F. Riggs, L. A. Journal of Experimental Psychology. Vol. 40, p. 687, 1950.
21. Smith, W. M., Warter, P. J. Eye Movement and Stimulus Movement: New Photoelectric Electromechanical System for Recording. Journal of the Optical Society of America. Vol. 50, p. 245, March 1950.
22. Cornsweet, T. N. New Technique for the Measurement of Small Eye Movements, Journal of the Optical Society of America. Vol. 48, p. 808, November 1958.
23. Polyak, S. The Retina. Univeristy of Chicago Press, Chicago, (1941).
24. Pirenne, M. H. L. Vision and the Eye. London Pilot Press, London, (1948).



5. Chemoreceptors

"The history of chemoreception is studded with man's attempts to explain and rationalize the mechanism involved in the activation of these sense organs. Yet today this is still one huge question mark. True, the work relating physical and chemical properties of compounds to their stimulating effectiveness may soon demonstrate those properties of the stimulus essential for activation of chemoreceptors, but the question of mechanism of activation will still remain unanswered. This answer must be sought through investigation of cellular changes in the sensory end organ itself, a difficult but challenging problem. (Ref. 1)"

The earliest animals arose from the sea and since they were sensitive to dissolved substances, changes in the concentration of these or other substances would produce a response that was always an avoiding action. This primitive common chemical sense differs from the higher developed senses of taste and smell in so far as it is stimulated by "foreign" injurious materials (ammonia, chlorine) and by naturally occurring substances (acids, alkalis, salts) only if their concentration departs widely from the normal value. Delicate receptor organs were developed for the senses of taste and smell. The sense of taste is much less delicate than the sense of smell (Ref. 2).

All the possible tastes are differentiated into four fundamental components: salt, acid, sweet and bitter. The mechanisms of taste stimulation of each will be discussed after a brief consideration of stimuli detection and the electrophysiology of response transmission. This is followed by a summary of the more widely held correlations of physical and chemical properties with taste. This section on the taste chemoreception concludes with a general discussion of sensitivity and some specific analogies to the other chemoreceptors.

The sense of odor is divided into the same main categories as the sense of taste. The brief introduction is followed by electrophysiological considerations of stimulus and response. The generally held physical and chemical correlations and theories of the mechanistic processes involved in odor stimulation are presented next. Because one of the most interesting facets of biosensors is their high sensitivity, the section on odor is concluded with a discussion of some recent work done on the absolute sensitivity of the sense of smell. One is amazed to



discover that the sense cells of the eye respond to one quantum of light with a macroscopic reaction and that the sensitivity of the ear and other organs of the labyrinth comes close to that of Brownian motion. Equally amazing are the experimental results which indicate that the olfactory cells respond to close to one odorous molecule. Because such low concentrations (only a few molecules) of stimulants appear to be adequate to evoke biological responses, it is presently thought that the stimulus only triggers the metabolic energy available in the cell (Ref. 3).

5.1 Taste

The receptor organs for taste are the taste buds. Gustatory papillae (the fungiform, foliate and circumvallate) which cover the surface of the tongue contain one or more taste buds, complete with several spindle shaped receptor cells, each with a hair like end projecting through the pore of the bud into the mouth cavity.

5.1.1 Detection of Stimuli and Transmission of Responses

In man, the sensation of taste is aroused when solutions of appropriate composition are in the mouth. The taste buds are the receptor organs for taste (Ref. 2). Just what happens after the contact has been made by the stimulating substance and the taste bud is not known. The penetration of the stimulating substances into the sense cells or their adsorption on the cell membranes are important events in stimulation which lead to the initiation of the nerve impulses (Ref. 1).

The taste receptor cell is only a generator of nerve impulses. Present at the periphery of the gustatory afferent system are: (1) a transducer, converting chemical changes to neural excitations (this is the taste cell) and (2) one or more centripetally conducting sensory fibers excited by action of the transducer (these are the gustatory afferents) (Ref. 4).

By employing electronic equipment it is possible to amplify and record nerve impulses in nerve fibers at points between the receptor cells and the brain. Most of these studies have been restricted to animals. Nerve fibers are very poor electrical conductors but they can transmit small current at the rate of 1-100 m. sec⁻¹ possible because the energy necessary for propagation is released pointwise



as the impulse proceeds. Nerve cells can be compared to batteries with a potential difference across the cell membrane. One theory is that the stimulus causes a depolarization of the nerve fiber ending and that a wave depolarization changes move along the nerve fiber. After 0.5 to 2 millise., each nerve section recharges and is ready to transmit another impulse. Impulses are recorded on an oscillograph. Electrical activity is observed when the receptor cell is resting but stimulation of the receptor site produces a change in frequency and amplitude of the nerve impulses. The potential difference across the semi-permeable cell membrane is due to different ion concentrations on either side. The extracellular fluid is higher in sodium ion concentration whereas the intracellular fluid is higher in potassium ions. Depolarization of the cell occurs when sodium ions are suddenly allowed to pass through the cell wall (Ref. 5).

5.1.2 Mechanisms of Taste

Electrophysiological methods have shown that most mammals possess receptors which respond to salt, acid, sugar and quinine solutions (Ref. 2). Some chemical relationship is expected between substances responsible for each of the four component tastes and the corresponding receptor organs (Ref. 2).

To be tasted a substance must normally be soluble in H_2O so that the taste stimulus could be a solid, liquid or gas provided that it dissolves somewhat upon contact with saliva. Some of the variables that determine the efficacy of taste stimuli are solubility, concentration, ionization ability, temperature, basic chemical composition and structure (Ref. 4). The mechanisms of sour, salt, sweet and bitter sensations will now be discussed.

Sour

The stimulus for the sour taste is an acid. Acids ionize in aqueous solution into an anion and cation of which the latter is the hydrogen or hydronium ion. The degree of sourness is related in a rough way to the degree of dissociation so that the sourness of solutions increases with the acidity of the solution which is inversely proportional to the pH of the solution. It has been shown that the taste of mineral or inorganic acids is not due to the anions but to the cations (Ref. 6). For example, acids such as HCl and H_2SO_4 cannot be distinguished by taste.



Organic acids and their anions have distinct tastes and may be more sour than expected from the pH of their solutions. Some acids are both sour and bitter such as picric acid, and some are sour and sweet as citric acid.

The addition of salts of weak acids depresses the ionization of the acids themselves so that in order to keep the same pH as would be present with the pure acid more acid would have to be added and this solution, now containing more acid and salt, tasted more sour than pure acid at same pH (Ref. 6). These experimental results indicate the influence of either the increased anion concentration or an ionic strength effect. In view of the results discussed later, the latter effect is the probable cause.

In order to arrive at a single, simple correlate for the sour taste it is necessary to eliminate the interaction between the sour stimuli and saliva, the latter serving as a buffer solution in the ensuing stimulation of the gustatory receptors. When the effect of saliva is minimized the results of threshold values for all acids of equal chemical combining capacity are uniform. The buffering action of saliva contributes to the discrepancies observed in acid detection tests but the minimization of the effect does not result in equal afferent nerve discharges at equal pH, equal normality or molarity.

There is some basic physiological mechanism which complicates the relation between sourness and acidity. The following are some of the suggestions that have been proposed:

1. Hydrogen ions might react with some substance on the receptor surface so that, as these ions in a solution of the organic acid were taken up, further dissociation would replace them;
2. Both the potential as well as the actual hydrogen ion concentration may be determiners of sourness.

A relationship has been noted between the sourness and swelling of gelatin in different acids. Sourness has also been attributed to the rate at which the acid penetrates the cell or intracellular spaces or to adsorption on the cell surface (Ref. 7).



Salt

The stimuli for the salty taste are soluble salts which dissolve in water to yield a cation and an anion both of which contribute to taste quality and stimulating efficiency. The typical salty taste is that of table salt (sodium chloride). However, not all salts give the salty taste. For example (Ref. 5):

SALTY

LiCl	LiBr	LiI	NaNO ₃
NaCl	NaBr	NaI	KNO ₃
NH ₄ Cl	NH ₄ Br		Na ₂ SO ₄
KCl			
RbCl			

SALTY AND BITTER

KBr	NH ₄ I
-----	-------------------

BITTER

CsCl	RbBr	KI	MgSO ₄
	CsBr	RbI	
		CsI	

SWEET

Pb(OAc) ₂
Be(OAc) ₂

There is a trend from salty to bitter taste with increasing molecular weight and also a trend in effectiveness of taste stimulation with ionic mobilities (Ref. 5).

One conclusion drawn was that the taste of a salt depends on the cation which determines the intensity and upon the anion which gives the character (Ref. 6). In a series of sodium salts, the quality of taste elicited will vary with the anion. The following anion series was obtained when the degree of saltiness of



various sodium salts was compared: $\text{SO}_4 > \text{Cl} > \text{Br} > \text{I} > \text{HCO}_3 > \text{NO}_3$. The following cation series is obtained if various chloride salts are compared:

$\text{NH}_4 > \text{K} > \text{Ca} > \text{Na} > \text{Li} > \text{Mg}$.

The thresholds for different salts have been variously reported to be equimolar for the cation, for halogen salts, inversely related to the molecular weight, and directly related to cation mobility (Ref. 7).

Certain trends and regularities were noted by employing the following taste equation in attempts to objectively specify the complexity of taste:

$N = xA + yB + zC + vD$ where x , y , z and v are molar concentrations and A stands for sodium chloride; B , quinine sulfate; C , fructose; D , potassium tartrate; and N is the molar concentration of the salt being matched (Ref. 7).

Thus, it appears that the saltiness of a particular compound can be expressed as a linear sum with the variable coefficient of the saltiness of these four particular compounds. The usefulness of this empirical correlation is questionable.

A theory of the mechanism of taste stimulation has been proposed (Ref. 8). Assuming a monomolecular reaction between the taste stimulus and some part of the taste cell, then:



where A is the stimulus; B , unfilled receptor sites; AB , filled receptor sites; C , concentration of stimulus; N , total number of available sites; Z , the number of sites filled at concentration of stimulus C .

Then the equilibrium constant may be written as:

$$K = \frac{(Z)}{(C)(N-Z)} \quad (2)$$

Assuming that the magnitude of response is proportional to the number of sites filled, and that the maximum response occurs at a high concentration of stimulus when all the sites are filled, then:

$$R = \alpha Z \quad \text{and} \quad R_s = \alpha N$$

R is the magnitude of response and R_s the magnitude of maximum response. Substituting the information into Equation (2), one arrives at the fundamental taste equation (Ref. 8).



$$\frac{C}{R} = \frac{C}{R_s} + \frac{1}{KR_s} \quad (3)$$

A plot of C/R against C (experimental data) yields a straight line which shows agreement of the data with the theory. This does not mean that the assumed model is a correct one, but the equilibrium constant can be found from the graph and this gives a measure of the binding force of the stimulus with the receptor. Equation (3) is similar to Langmuir's adsorption isotherm, to many enzyme-substrate equations and to the equation expressing the binding of ions by proteins. The small equilibrium constants found for various sodium salts are in agreement with the concept that the stimulus is adsorbed to the cell surface and that the reaction is not an enzymatic one. The equilibrium constants are of similar magnitudes to those in the adsorption of salts to proteins.

CALCULATED EQUILIBRIUM CONSTANTS (REF. 8)

<u>Sodium Salts</u>	<u>Equilibrium Constant K liter/mole</u> <u>(Ref. 9)</u>
sodium chloride	9.8
sodium formate	9.0
sodium acetate	8.55
sodium propionate	7.58
sodium butyrate	7.72

When the temperature is raised from 20 to 30°C the magnitude of the taste receptor response to salts does not change, therefore, $\Delta H = 0$, indicating that this spontaneous process is physical and not enzymatic. Thus, $T\Delta S > \Delta H$ in the equation:

$$\Delta F = \Delta H - T\Delta S \quad (4)$$

From the equilibrium constant, ΔF can be calculated

$$\Delta F = -RT \ln K \quad (5)$$

so that the change in free energy, ΔF , is negative and thus the change in entropy,



ΔS , is positive. This indicates that the water of hydration of the ion (for example, the water of hydration normally surrounding the sodium ion) decreases when the ion is adsorbed to the receptor surface or else that the receptor molecule changes its spatial configuration slightly when the ions are adsorbed to it, or both events could take place simultaneously. This is similar to the small positive changes in entropy when ions are bound to proteins (Ref. 7).

The binding may also be studied by changing the pH of the solution. The magnitude of the response to 0.1 M NaCl does not change appreciably when the pH is varied from 3 to 11 implying that most carboxyl units of proteins are not involved in stimulation, however, other groups such as phosphate groups may be involved, but the reasoning is not completely evident. According to the present theory of taste, the receptor molecule may be regarded as a polyelectrolyte containing a large number of charged side chains that make up the sites (Ref. 8). (The carboxyl units should be H-bonded to the amino H of the neighboring protein chain.)

Beidler (Ref. 9) suggested a possible mechanism of receptor stimulation to explain how the adsorption of a chemical stimulus on the surface of a receptor can cause an increase in the electrical activity of the nerve innervating that receptor. The taste bud cells like most receptors are negatively charged relative to their exterior environment caused by different concentrations of ions in the cells and outside the cells. When an electrolytic or non-electrolytic stimulus is adsorbed on the receptor site of the taste cell then a slight change in spatial configuration of the receptor molecule may occur so that a hole is formed that is large enough for certain ions (most likely potassium) to escape from the inside to the outside of the cell thus decreasing the potential across the receptor membrane. A spread of local depolarization over the cell surface may by chemical or electrical means stimulate the innervating nerve so that the frequency of nerve impulses generated is proportional to the magnitude of receptor depolarization.

Some substances such as sodium chloride and sugar taste salty and sweet over large ranges of concentrations, whereas a large number of substances possess mixed tastes. Ammonium chloride tastes salty at low concentrations, both bitter and salty at medium concentrations, and bitter at high concentrations. Different receptor sites with different equilibrium constants are predicted to be involved in these responses. The sites involved at low concentrations usually



have a higher equilibrium constant for the stimulus-site reaction and a lower level of saturated response, R_s , than do the sites activated at higher stimulus concentrations. The relative number of either type of site may vary from one receptor to another. Since, ammonium chloride is bitter at high concentration, it is expected that a greater number of sites with low K values are found in those receptors that also have an abundance of sites reactive to alkaloids and other bitter substances (Ref. 9).

The present theory of taste, i. e. , considering the receptor molecule as a polyelectrolyte containing charged side chains that make up the receptor sites, was developed for the interaction of salts with a receptor, but it has also been used to describe acid, sugar and bitter stimulation. Even though the type of binding of sugars is different from salts, a monomolecular reaction of adsorption is held. Acids stimulate by means of their hydrogen ions although the surface adsorption depends on the amount of un-ionized acid in the solution and the ionic strength. A strong acid is a more efficient stimulant than a weak acid of equimolar concentration. At the same pH value, the weak acid is a more effective stimulant. By making a buffer solution of a weak acid the sourness does not decrease to the extent that would be expected from the change in pH. This is accounted for by the increase in ionic strength which tends to increase the amount of acid adsorbed at a given pH (Ref. 8).

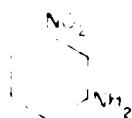
Sweet

It is mainly organic compounds except for some salts of lead and beryllium that give rise to the sweet taste. Polyhydroxy and polyhalogenated aliphatics are usually sweet, e. g. , glycerol, sugars, and chloroform. Other stimuli are aldehydes, ketones, amides, esters, amino acids, and sulfonic acids. Many attempts have been made to relate chemical structure and taste. Several limited correlations have been made but no widely applicable theory has evolved. Oertly and Myers listed a number of sweet-producing molecular arrangements and postulated that to be sweet, a substance must contain a glucophore and an 'auxogluc'. Saccharin is an important exception in this classification (Refs. 5 and 7).

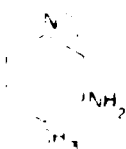


In an homologous series, the taste of the members often changes from sweet to bitter with an increase in molecular weight. Taste often disappears with the higher members of the series because they become water insoluble (Ref. 7).

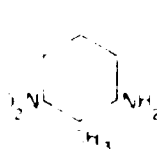
Small changes in the structural arrangement of homologues produces striking changes in taste. Considering the homologues of m-nitro aniline which is sweet, only 2-nitro-p-toluidine is sweet.



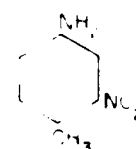
SWEET



SWEET



TASTELESS

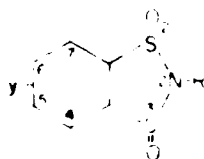


VERY SLIGHTLY BITTER

It has been concluded by some investigators that after studying the relation of taste and structure that it was impossible to derive any law for aromatic compounds and that sweet taste depends not on any single factor, such as a certain taste-producing group, but on the entire chemical complex of the particular compound studied (Ref. 10).

One of the better-known sweetening agents is saccharin. In 1914, Cohn summarized some information about saccharin and its related compounds and concluded that sweet taste is lost if the sulfimide ring is opened, or if the imide hydrogen is replaced; and that substitution in the benzene nucleus reduces the sweet, and introduces a bitter taste (Ref. 10).

In 1961 Hamor reported correlations of taste with chemical structure of approximately 80 derivatives of saccharin (Ref. 10).



GENERAL FORMULA FOR SACCCHARIN DERIVATIVE



SACCCHARIN

Some generalizations that may be drawn from the results are given below.

Replacement of the imide hydrogen with another chemical group gave, in almost every instance, a tasteless compound. Both sweet and bitter compounds were converted to tasteless substances by this substitution in the 2-position.

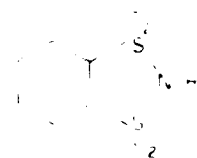


There were three exceptions where the bitter taste that resulted may be due to the basicity of the relatively high molecular-weight tertiary amino group. These results appear to lend evidence to the belief that sweet taste is attributed to the anion. It is possible that the sense of taste resembles other types of physiological responses and is subject to the "lock and key" concept of biological activity. Then substitution on the imide nitrogen may affect the specific molecular shape which must fit the receptor site so that the 2-substituted compounds are not complementary to the taste receptors, and display no taste. Yet some workers have reported 2-substituted saccharins which do not ionize or hydrolyze but have a sweet taste (Ref. 10).

Another possible explanation for the lack of taste of the 2-substituted saccharins may be that isomerization of the lactam form to the lactim form is necessary for sweet (and bitter) taste.



LACTIM FORM



o-BENZENEDISULFONIMIDE

If groups are substituted in the 2-position, then there is no possibility of forming the lactim form. If an alkyl group is substituted for the imide hydrogen in o-benzenedisulfonimide, the taste changes from sweet to tasteless which is analogous to the results of 2-substitution in saccharin.

If electron withdrawing groups such as the nitro group are substituted in the benzene ring then bitter substances result. If the nitro group is reduced to the electron donating amino group then sweet compounds result.

In addition to the inductive effect, resonance may contribute to taste. The 4- and 6- nitro saccharins which are ortho and para, respectively, to the carbonyl group, contribute to the sweet taste whereas the 5- and 7- nitro saccharins, meta to the carbonyl group, give only a bitter taste. Perhaps, the resonance contributions of these groups to the relative acidity of the molecule, or to the relative ease of formation of the lactim form, is responsible for the differences in taste of the saccharin derivatives.



Substitution in the 3-position produces tasteless compounds as does substitution in the 2-position. If the heterocyclic ring is opened, tasteless compounds are produced (Ref. 10).

Stereoisomerism is significant in taste as well as other physiological systems. The dextro form of asparagine is sweet, the levo form tasteless. Solutions of α -D-glucose are sweeter than β -D-glucose which predominates in solution after mutarotation has occurred.

The degree of sweetness of various sugars was measured and no relation was observed between the number of C atoms, OH groups, or the molecular arrangement or spatial configuration of the compounds (Ref. 6).

Bitter

Bitter, like sweet, is a taste sensation that is elicited by members of many chemical classes and is found associated with sweet and other taste qualities. Increasing the molecular weight of inorganic salts is associated with increasing bitterness. An increase in the length of the C-chain in organic molecules sometimes produces a change from sweet to bitter. Many sweet substances such as saccharin have a bitter after taste. This multiple taste quality is noticeable when the stimulus moves from the front of the tongue to the back where bitter sensitivity is the greatest (Ref. 7).

The complex nitrogenous compounds, the alkaloids, are the best known class of bitter compounds, i. e., quinine, caffeine, strychnine, and nicotine. Most nitro compounds are bitter (picric acid). The following groups are most often associated with bitter taste: $(\text{NO}_2) > 2$, $\equiv \text{N}$, $\equiv \text{N} =$, $-\text{SH}$, $-\text{S}-$, $-\text{S}-\text{S}-$, and $-\text{CS}-$ (Ref. 7).

The importance of structure and chemical grouping is demonstrated by the phenomenon of "taste blindness" which is an insensitivity to substances having the $\text{>NC} \begin{smallmatrix} \text{S} \\ | \end{smallmatrix}$ - group whereas $\begin{smallmatrix} \text{H} \\ | \\ \text{N} \end{smallmatrix} - \text{C} \begin{smallmatrix} \text{O} \\ || \end{smallmatrix}$ is the familiar "peptide linkage." Phenylthiocarbamide (PTC) is used as a test for "taste blindness". This inability to taste PTC is genetic and inherited as a Mendelian recessive characteristic. These taste defects were found in anthropoid apes as well as man.



S
 $\parallel$

The inability to taste $>\text{NC}$ - may be overridden by other chemical groups as in thiourea, NH_2CSNH_2 , which is sour to all persons. Taste blindness is not correlated with sensitivity for other bitter stimuli or other taste qualities, which suggest a high degree of specificity for a particular chemical linkage connected with some feature of the receptor mechanism. Attempts to link taste blindness with insolubility of PTC in saliva were unsatisfactory. Because bitter and sweet sensitivity are often associated with similar type stimuli, and both are inactivated by drugs or narcotic agents, the proposal has been made that both depend upon the action of a single receptor mechanism (Ref. 7).

Of all the taste mechanisms, that underlying the bitter sensation is least well known.

5.1.3 Physical Property Correlations of Taste

Certain physical properties of stimulating compounds have been studied to determine their effect on taste. For example, the surface tension of solutions of substances could affect taste by their ability to penetrate the taste bud or alter the permeability of taste bud cells. Since all the compounds that effect a sweet taste have chemical groups that may hydrogen bond with water, it is likely that the degree of association with water might be a factor (Ref. 5).

Since enzymes take part in all biological processes it is possible that they might be involved in the excitation stage of taste stimulation. If this is a redox enzyme system, then the ease of oxidation or reduction of a compound will bear on its taste ability. In attempting to define generalizations between taste and the physical properties of stimulants, only limited correlations result. For example, the relative sweetness of the four monosaccharides fructose, glucose, galactose and mannose decrease in that order with their half-wave reduction potentials. When more compounds and other physical properties listed above are included the relationship fails (Ref. 5).

Other theories have attempted to explain taste from the standpoint of cell physiology. The cell membrane has been regarded as a water balanced emulsion system of water-in-oil and oil-in-water types and that taste resulted from the adsorption and surface tension action of substances in this system and that the



subsequent changes in cell membrane permeability were due to changes in the phase ratio of the emulsion types. Then it was proposed that the solubility and permeability in cellular solvents of compounds were controlled by their structure which determined their thermodynamic activity and thereby their tastes. The thermodynamic activity was taken as the ratio of their active concentration to their solubility. These approaches have not been successful.

The process of adsorption has received quantitative support as being involved in the critical step of taste stimulation. For example, the sweetness and concentration of sugars are related by the following equation

$$S = KC^m \quad (5)$$

where S is sweetness; C is concentration, K and m are constants. This is the Freundlich equation used to express adsorption from solution on solid surfaces. It was found in studies made on the taste nerve fiber of the rat that as the concentration of a salt stimulus increases, the response rises asymptotically to a maximum level.

It has been found that geometric and optical isomers have different tastes. Anti and syn isomers of oximes have different tastes. D- α amino acids are usually sweet whereas the natural L isomers are flat to bitter. For example:

Isoleucine: d (-) sweet	l (+) bitter
Valine: d (-) sweet	l (+) flat to bitter
Leucine: d (+) very sweet	l (-) flat to bitter
Histidine: d (+) sweet	l (-) flat to bitter

Consequently, the spatial configuration of a compound may have a critical effect upon its taste. It has been believed that adsorption on some stereospecific surface of the taste organ is the initial or significant act in the sequence of steps involved in taste stimulation. The stereospecific surface may be that of an enzyme (Ref 5). Work on the chemoreceptors of the blowfly indicates a possible role of stereospecific interactions in taste stimulations.



5.1.4 Chemical Constitution and Taste

Moncrieff has studied the evidence between taste and chemical constitution and compiled the following list of general relations (Ref. 6).

"

1. The sour taste is characteristic of acids and acid salts. The primary cause is the hydrogen ion.
2. The degree of sourness runs parallel with the hydrogen-ion concentration, but is anomalously high when the acid has lipoid-solvent properties.
3. Not all acids are sour. A bitter or sweet taste may overwhelm the sourness, e. g. , picric acid is bitter and salicylic acid is sweet. Nearly all acids containing the group -COOH are sour.
4. The saline taste is characteristic of salts. They ionize in the saliva.
5. The anion, the cation, complex ions and the undissociated salt all contribute to the taste.
6. Salts of low and medium molecular weight are salty in taste, but heavy salts such as iodides or salts of caesium are bitter. Heavy cation and heavy anion certainly give bitterness, but if one ion is heavy and the other light the result cannot be predicted.
7. Beryllium salts and some lead salts are sweet. Magnesium and ammonium salts are bitter. Salts of heavy metals such as mercury usually have a metallic taste.
8. Substances may have more than one taste, e. g. p-chloroaccharin is sweet and bitter and pyridine 2,5 dicarboxylic acid is sweet, bitter and sour.
9. Unlike the salt and sour tastes, the sweet and bitter are not confined to single chemical groups but are found in particularly every class.

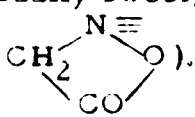


10. Polyhydroxy compounds are usually sweet, e. g. glycols, sugars (Cohn's sapophore, $(OH)_x$).
11. α -amino acids are usually sweet (Cohn's dulcigen group $\begin{array}{c} \diagup \\ \text{C} \text{---} \text{NH}_2 \\ \diagdown \\ \text{COOH} \end{array}$). β - and γ -amino acids are not usually sweet, particularly the latter. The closer together the amino and carboxylic groups the greater the sweetness.
12. On ascending a homologous series the taste frequently changes from sweet to bitter.
13. The influence of any particular group is less in a large molecule than in a small one.
14. On ascending a homologous series taste eventually disappears, simultaneously with solubility.
15. The presence of three nitro groups in a molecule gives a bitter taste. Two nitro groups usually give a bitter taste. One nitro-group often gives a sweet taste, e. g. nitrobenzene.
16. Alkylation of an amine group usually gives a sweet taste.
17. Alkylation of an amide group often gives a sweet taste, e. g. α -dimethylurea.
18. Alkylation of an imide group destroys sweetness, e. g. N-methyl saccharin is tasteless.
19. Etherification of a hydroxyl group often abolishes a sweet taste, e. g. glycol monomethyl ether is not sweet.
20. Many of the changes which increase odour, decrease taste and vice versa. Lipoid-solubility is primarily important for odour, water-solubility for taste. Strongly sapid substances are often odourless.
21. Alkylation of the benzene ring often gives a sweet taste, e. g. p -tolunitrile is sweet but not benzonitrile.



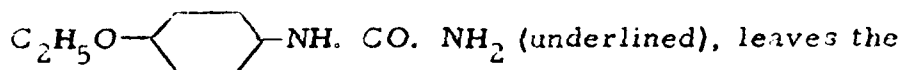
22. The entry of a phenyl group usually causes a bitter taste, irrespective of whether it is attached to carbon, oxygen or nitrogen. This may partly be due to the increase in molecular weight so caused.
23. Chain branching reduces the sweet taste and introduces a bitter or a burning pungent taste.
24. Introduction of a methoxy group frequently gives a sweet taste, e. g. anisonitrile.
25. Esters with fruity odours usually have sweet tastes, e. g. ethyl butyrate. Esters of inorganic acids are mostly sweet.
26. Esters of aromatic acids, sulphinic acids, as well as lactones and urethanes are bitter.
27. Esters of resorcinol and allied substances are usually sweet, but no general relation exists between the taste of an alcohol and the ester derived from it.
28. Unsaturation often brings pungency, e. g. allyl nitrite, probably an instance of the common chemical effect overpowering the taste.
29. Nitroparaffins are sweet, but unsaturation and chain branching give pungency and a caustic taste.
30. The amino group often brings a sweet taste, especially if the positive effect of the amine group is balanced by the proximity of a negative group, such as carboxyl or nitro.
31. Acylation of the amino group reduces or destroys the sweet taste.
32. Free bases are bitter, particularly alkaloids and tertiary ammonium bases (Cohn's groups $N \equiv$ and $=N \equiv$).



33. The simplest amides are bitter. Amides of dibasic acids are tasteless. Many amides are sweet.
34. Dihydrazides, where the two hydrazide groups are close together, have sweet tastes (oxalic exceptional), but not where far apart.
35. For ureas to have a sweet taste it is necessary for one of the NH_2 groups to be unsubstituted, hence symmetrical ureas are not sweet.
36. Neutral nitrogen compounds are often sweet, e. g. α -amino acids, dihydrazides, ureas, oximes (usually sweet).
37. Betaines are bitter (Cohn's sapophore ).
38. In the aliphatic series heavily halogenated bodies are often sweet, e. g. chloroform.
39. In the aromatic series, a halogen in the nucleus introduces a bitter note. The bitter effect is least with fluorine and greatest with iodine.
40. Sulphur in aliphatic compounds induces a bitter taste. Mercaptans, sulphides, disulphides, thioamides, and thioureas are bitter (Cohn's groups $-\text{SH}$, $-\text{S}-$, $-\text{S}-\text{S}-$, $-\text{CS}-$). The sulphonic acids, $-\text{SO}_3\text{H}$, are sour or bitter.
41. The effect of a sulphur atom as a ring member is quite different. The ring functions as a whole, e. g. saccharin has heterocyclic sulphur.
42. The effect of a sulphur-containing group substituted in the benzene ring is to reduce the sweet taste and throw it towards bitter.
43. The sweetness of saccharin is due to the anion. Its alkali and alkaline earth salts are sweet, but saccharin salts of heavy metals are astringent.



44. The sweet taste of saccharin is lost if the sulphimide ring is opened or if the imide hydrogen is replaced.
45. Substitution in the benzene ring of saccharin reduces the sweet and gives a bitter taste.
46. p-alkoxy phenyl ureas are sweet, e.g. dulcin. If the terminal NH_2 group is substituted taste is lost (cf. 35).
47. Doubling of the phenyl group inhibits the taste effect in both saccharin and dulcin (cf. 22).
48. Small changes in the ethoxy group of dulcin to methoxy, hydroxy, chlorethoxy, etc., leave the sweet taste still there, but larger groups like propoxy and acid groups destroy the sweet taste.
49. Substitution of the hydrogen on the already substituted nitrogen,



sweet taste but gives a bitter after-taste.

50. Some aldehydes are sweet, but no ketones. Semicarbazones of sweet aldehydes are less sweet, and phenylhydrazones are not sweet. Oximes often are sweet.
51. Certain sapophore groups are often associated with sweet tastes and others with bitter tastes. The most important are
 sweet: $(\text{OH})_n$, $-\text{CH}(\text{NH}_2)\text{COOH}$, $-\text{C} \begin{array}{l} \text{H}_{3-n} \\ \text{Cl}_n \end{array}$, $-\text{CH}_2\text{ONO}_2$.
 bitter: $(\text{NO}_2)_2$, $\equiv \text{N}$, $=\text{N}\equiv$, $-\text{SH}$, $-\text{S}-$, $-\text{S}-\text{S}-$, $-\text{CS}-$.
52. Not only the groups present, but also their arrangement determine taste.



53. Homologues usually have similar tastes if close to each other in a series, e. g. formamide, acetamide, propionamide, all bitter.
54. Isomers with considerably different constitutions have no taste relationship.
55. Close isomers may have different tastes, e. g. leucine sweet, isoleucine sweet and pungent, but there is usually some resemblance.
56. Substitution in the benzene nucleus gives isomers with unpredictable tastes. Often ortho and para are similar but different from meta.
57. Stereoisomers may have different tastes. The dextro rotatory isomer usually has the sweeter taste, e. g. d-valin sweet, l-valin tasteless.
58. Taste depends not only on the groups present but also on their arrangement in space. "

5.1.5 Sensitivity

The sensitivity of receptors is expressed by the threshold concentration necessary to elicit a minimum response. This sensitivity is related to the binding strength of the stimulus to the receptor site if one considers the fundamental taste equation, $\frac{C_T}{R_T} = \frac{C}{R_s} + \frac{1}{KR_s}$ because if the concentration of the stimulus is decreased, the value of C/R_s approaches zero, and the value of the magnitude of response at threshold concentration is

$$R_T = C_T K R_s \quad (7)$$

It can be seen that the threshold concentration, C_T , not only depends upon the strength with which the stimulus is attached to the receptor site as measured by the equilibrium constant, K , but also upon the maximum number of sites available for that particular stimulus, R_s . Thus, by measuring thresholds one only obtains



a relative measure of the binding strength of a series of substances in those cases where it is certain that the maximum number of available sites for these substances is constant. Since R_g may vary with the type of stimulus, then the threshold concentration alone is not a good measure of the binding strength of the stimulus to the receptor (Ref. 9).

One can determine if there is a separate single fiber carrying information about sweet to the brain, and separate ones for bitter, sour and acid by dissecting free a single taste fiber and recording its neural activity. The change in response is similar to that obtained from all other sensory cells; that the frequency of the nerve impulse increases with the intensity of the stimulus to the receptor. If a number of taste nerve fibers is studied, each fiber is observed to differ slightly in sensitivity from its neighbor, and most of the fibers respond to two or more of the stimuli representing different taste qualities. Thus, the complex pattern of activity from a number of single taste nerve fibers is thought to be necessary for quality discrimination (Ref. 9).

(A receptor may respond to a stimulus within 30-50 millisecc. after it is applied to the surface of the tongue.)

Weber first noticed that the smallest difference in the weight of two objects bears a constant relation to the weight of the objects. Weber's Law may be expressed:

$$\Delta I/I = C \quad (8)$$

Where I is intensity of stimulus, ΔI is the smallest noticeable change in intensity and C is a constant. Fechner developed this idea. He postulated that each discriminable step of stimulus intensity corresponds to a unit increase in sensation, then

$$\Delta I/I = k \Delta s \quad (9)$$

where ΔS is the increase in sensation. From this

$$dS/dI = 1/kI \text{ and } S = a \log I + b \quad (10)$$

This relationship indicates that sensation is proportional to the logarithm of the stimulus intensity (Ref. 11).



Experimental work on taste and smell indicates that the stimulus or concentration of the sapid or odorous substance must be increased approximately 30% to produce a noticeable difference in sensation, so that for the chemical senses $\Delta I/I$ has the high value of 0.3, in vision 0.01, in hearing 0.1, in pressure of the skin 0.05, and in **perception** of weight differences 0.025 (Ref. 6).

The minimum stimulus refers to that quantity of matter which just evokes sensation. Ethyl alcohol, C_2H_5OH , is a compound which evokes sensations of taste, odor and the general chemical sense and it offers an opportunity to compare the sensibilities of the three senses. The minimum stimulus for the three is as follows:

0.44% wt/wt in air for smell

14% wt/wt in water for taste

25% wt/wt in water for general chemical sensation (Ref. 6).

The threshold concentration for ethyl mercaptan, C_2H_5SH , is $0.3 \times 10^{-3}\%$ wt/wt which is 100 million times more delicate **than that for ethyl alcohol**.

If one compares strychnine hydrochloride, one of the most powerful sapid substances with mercaptan, one of the most powerful odorous substances, it can be concluded that odor is 10,000 times as sensitive as taste.

5.2 Odor

The olfactory mucosa lies in the dorsal and posterior part of the nasal cavity. It is a pseudostratified columnar epithelium with no distinct cilia. The olfactory receptors are hair cells which are bipolar and oval. **Electron microscopy** indicates that there are up to 1,000 hairs per cell. Each hair is 1 to 2μ long and 0.1μ in diameter. This increases the surface area of the receptor cell greatly. The secretions of the mucous glands bathe the nasal cavity in a liquid sheath which is in a constant state of motion toward the nasopharynx. This sheath is considered to be very important in conveying odorous substances to the receptor cell since solubility (water and lipid) in the mucosa appears to be related to the odorous properties (Ref. 12).



5.2.1 Electrophysiological Considerations

The electrical responses of the frog's olfactory mucosa were successfully recorded when odorous air was blown into the nasal cavity evoking a slow negative monophasic potential. The amplitude of the response is proportional to the logarithm of the stimulus intensity. The shape and time course of the response is related to the strength of the stimulus. With an increase of odor intensity in the stimulating air, the potential rises at a faster rate, the crest of the response broadens and the decay time lengthens. It has been found that during continuous stimulation the evoked response in the olfactory epithelium declines from the initial peak to a lower level which continues throughout stimulation (Ref. 12).

The changes occurring in the olfactory nerve where odors stimulate the receptor cells have been studied but they are not as well known as the analogous events of the other senses because of the inaccessibility of the fine, unmyelinated, weak nerve threads. Adrian has performed the most extensive experiments employing fine wire electrodes and inserting them into various portions of the olfactory bulb. Bursts of action potentials resulted from injections of odor-laden air. These electrical activity records represented the summed effect of many primary neurons (Ref. 4).

5.2.2 Mechanisms of Odor

Some of the essential processes involved in olfactory stimulation will be discussed here.

As yet, no complete or inclusive picture has been advanced to categorize all odorous substances because many are totally unrelated physically and chemically.

Hill and Carothers observed a relationship between the number of atoms in certain macrocyclic ring hydrocarbon compounds and their odors. A cedar-type odor was possessed by 13-atom compounds, a musk-type odor by those with 14, 15 or 16 atoms, and a civet-type odor with 17 or 18 atom compounds. They suggest that the number of atoms in the ring, rather than the identity of the reactive groups is the significant factor within wide limits. Moncrieff suggests that odorous substances are volatile and soluble (water and lipid) in the tissues of the



olfactory mucosa. The adsorptive properties of the receptors are important to processes of excitation. Because only seven elements are odorous (F, Cl, Br, I, O₂ as O₃, P and As) and six of them occupy the lowest places in the electrochemical series, Moncrieff correlates the disposition of an element in the electrochemical series with odorous properties. Substitution groups in organic compounds determine the strength and quality of odor. Legge has hypothesized that odorous substances may react with groups on the protein and lipoprotein film surface leading to the rupture of the protein bonds and a great increase in their area. The -S-S- bond is very important in maintaining protein structure, and rupture of these bonds by mercaptans might explain why these substances are so sensitive (Ref. 11).

Adrian has related the properties of electrical discharges with four groups of substances. Acetone is a member of a group including amyl and ethyl acetate. Benzene belongs to an aromatic hydrocarbon group, octane to a group of paraffin hydrocarbons and heavy oils; and dicyclopentadiene to a group including the terpenes (Ref. 11).

The ability of metals to adsorb and retain odors was tested by Deininger and Sullivan who found that many metal surfaces pick up, modify and distort odors (Ref. 12).

Many enzyme theories have been proposed. Alexander suggested that odor-producing substances affect the catalyst balance of the receptor cells. Kistiakowsky elaborated this theory and hypothesized that the change in concentration of the products of the reaction cause excitation. Sumner criticized this theory and stated that odorous substances in the concentrations needed to evoke responses would not affect the known enzyme systems and that this theory needed new and unusual enzymes (Ref. 12). No one has explained how a change in concentration caused by inactivation of an enzyme could stimulate olfactory nerves.

Goldwasser suggests that the energy needed to stimulate olfactory receptors may come from Pauling's electrochemical energy source derived from the change of bonding angles within a molecule when it goes into solution (Ref. 12). Many theories relating electromagnetic radiation or molecular vibration to olfaction have been proposed. There is no experimental foundation for the concept that the essential properties of odor result from radiations inherent in molecular behavior. It has been shown that certain substances have the same odor but different spectra.



whereas other substance who have identical spectra smell differently such as some d- and l-isomer. Electrical responses were not recorded when the olfactory mucosa was covered with a thin plastic membrane which allowed the passage of infrared radiation but prevented physical contact by the stimulating particles (Ref. 12).

Moncrieff (Ref. 13) has devoted a chapter in his book to the "Physical Properties of Odorous Materials", wherein he discussed work done and theories of odorous substances in relation to vapor pressure, solubility, infrared absorption, ultra-violet absorption, ultra-violet irradiation, refractive index, Raman effect, spray electricity, diffusion, diamagnetism, dilution and adsorption. He has also accumulated much information relating chemical constitution and odor. After examining the literature he selected the following general principles stating that "they are open to criticism, but the subject is in an early stage of development, and if they are criticized and replaced by a better lot they will have served their purpose (Ref. 13).

1. Compounds of different constitutions may have similar odors, e. g. camphor, benzyl alcohol, direne.
2. Compounds of very similar constitution may have different odors. If, however, the constitutional differences are slight, odor differences are generally correspondingly slight.
3. No element which occurs free in nature is odorous under normal conditions, but elementary arsenic is odorous if heated and vaporized.
4. Seven of the elements are odorous, viz. fluorine, chlorine, bromine, iodine, phosphorus, arsenic, oxygen (as ozone).
5. Polymerization reduces or destroys odor whether in elements, e. g. red phosphorus, or in compounds, e. g. glycols.
6. The elements give only two kinds of smell - "halogen" and "garlic" - which are possibly fundamental or pure odors.



7. Ability to function in a high state of valency is productive of odor. (The halogens are in Group 7, oxygen in Group 6, and phosphorus and arsenic in Group 5 of the Periodic Classification.)
8. Electronegative or non-metallic character is allied closely with odorant properties. Elements at the bottom of the electrochemical series are odor-producing.
9. Electropositive elements suppress odor formation. (Those in Groups 1 and 2 or at the top of the electrochemical series.)
10. Sulphur, selenium and tellurium compounds often have vile smells whilst their oxygen analogues are odorless.
11. Compounds of phosphorus, bismuth and arsenic often have garlicky odors.
12. Unsaturation enhances odor but does not initiate it. (The paraffins are odorous.)
13. In the paraffins, straight and branched-chain isomers have similar odors, and the position of the double bond (if any) is unimportant.
14. In a homologous series the odor will rise to a maximum as we ascend the series and will then fall off owing to decreased volatility.
15. Unsaturation often introduces an irritant note to the odor, particularly if close to a polar group, e. g. aliphatic aldehydes and acids.
16. A tertiary carbon atom will frequently induce a **camphoraceous** odor.
17. The osmophoric influence of the phenyl group is strong. It overcomes that of alkyl ether groups and also of the amino group.



18. The introduction of a hydroxyl group frequently depresses or abolishes odor, e.g. aliphatic acids.
19. Aromatic acids are usually odorless, whereas most aliphatic acids are odorous.
20. Fully reduced aromatic bodies have odors similar to those of fatty substances, e.g. cyclohexane, hexahydrobenzoic acid.
21. Fully reduced heterocyclic bodies also develop aliphatic odors, e.g. piperidine.
22. Esters have fragrant fruity odors. This includes esters of weak inorganic acids such as boric.
23. Strong odor is often found accompanied by volatility and chemical reactivity, e.g. aldehydes. Chemical reactivity and unsaturation run parallel to odor.
24. Esters of diacids have good odors, but those of dihydric alcohols have weak odors.
25. Ketones generally have pleasant odors.
26. Whereas introduction of a hydroxyl group frequently destroys odor, the etherification of this group usually restores the odor.
27. Lactones have fragrant ester-like odors.
28. The alkyl halides have sweet smells, which become heavier in type as we pass from chlorine to bromine and to iodine. Two or three halogen atoms have an additive olfactory effect.
29. When a halogen is already a substituent, a hydroxyl group behaves abnormally and increases the odor.
30. The odor of amines is more ammoniacal when concentrated than when dilute, and the fishy odor is stronger in dilute than in concentrated solutions.



31. Nitrogen compounds frequently have an "animal" odor. Compounds not containing nitrogen do not ordinarily have an "animal" odor.
32. An oxygen linkage is frequently associated with a pleasant odor, e. g. esters, lactones, nitrates.
33. Many nitriles have a bitter-almonds odor.
34. Compounds in which an element functions at a valency lower than its maximum usually have offensive odors, e. g. hydrogen sulphide and the isonitriles.
35. Many sulphur compounds, particularly where the sulphur is divalent, have offensive odors, e. g. mercaptans.
36. Sulphur and nitrogen and oxygen atoms in a ring do not play their usual olfactory role. Their influence is subordinate to that of the ring, e. g. thiophen, furan.
37. If a heterocyclic body is completely reduced, the sulphur or nitrogen or oxygen atoms regain their normal olfactory effect.
38. Whereas unsaturation is often a contributory factor to odor, ring unsaturation appears to have a negative olfactory action.
39. The main factor in determining odor is the architectural type of the molecule.
40. The osmophoric groups play a part subordinate to that of the architectural style. In small molecules this part is greater than in large molecules.
41. The olfactory offensive and delicious are often closely related.
42. Some substances change their odor on dilution, e. g. indole, amines.
43. In ring compounds, the number of ring members often determines the odor:



5-6 members give bitter almonds and menthol odor.
6-9 members give transitional odor.
9-12 members give camphor or mint odors.
13 members give woody or cedar odors.
14-16 members give musk and peach odors.
17-18 members give civet odors.
More than 18 give faint or no odor.

The odour of compounds with rings from 9-20 members is largely independent of the substituent groups.

44. For a musk odor in macrocyclic compounds the basic ring structure must be between 14 and 19 atoms, and at least one carbonyl or imine group is necessary. Musk odors are not confined to macrocyclic compounds.
45. One heterocyclic atom increases the odor of macrocyclic compounds, two reduce it.
46. In the ionones, the position of the double bond in the ring is unimportant, but unsaturation in the side-chain is essential for the violet odor.
47. Arsenic compounds are usually garlicky or ill-smelling (cacodyl), but not if the arsenic is a heterocyclic member of a ring.
48. In a mixed ether-ester compound, the ester group has a dominant olfactory effect.
49. Branching of a chain usually enhances the odor.
50. There is no relation between the odors of isomers with important structural differences.
51. Structurally similar isomers usually have similar odors. As the difference in molecular structure or architectural style of isomers increases, so do odor differences.



52. The position of a side-chain influences the odor. The nearer the side-chain to a polar group, the stronger the odor, e. g. α -substituted lactones have stronger odors than γ -substituted.
53. Position isomerism in the benzene ring has a marked influence on odor, e. g. o-biphenyl ethers are odorous, para are odorless.
54. The 1. 3. 4 arrangement in the benzene ring usually leads to a pleasant odor, but not if the substituent groups are heavy.
55. Para-substitution often introduces a soft anise note.
56. Meta-substitution usually enhances the power and pungency of an odor.
57. Ortho-substitution often gives a floral odor.
58. With equal substituents the para influence overwhelms the meta (or ortho) - if both are present.
59. Where methyl and chlorine occur as substituents in a benzene ring the latter has the greater influence.
60. The odor of a substituted aromatic compound depends more on the position than on the nature of the substituent groups.
61. Stereoisomers may have different odors. The differences are not usually very great.
62. Finally, the general conclusion is that odor is determined by the architectural arrangement of the molecule. Osmophoric groups (sic) have a secondary influence, only in so far as they modify the general arrangement. The great difficulty is, of course, the subjective nature of odor classification. Once this difficulty is removed and it is possible to define exactly the quality of an odor, then logic and precision will reduce chaos to order, and the relation between odor and chemical constitution will be disclosed to us with mathematical exactness."



A history of "Theories of Odor" was presented by Moncrieff. He believes that, "all contain some of the truth, but none all the truth... together they may embrace most of the truth." The theories contain the following main factors: volatility, solubility, reactivity (olfaction probably includes some chemical reaction), and vibrations (intramolecular).

More recent work has placed emphasis on the stereospecificity of chemoreceptors insofar as the senses of taste and odor resemble other physiological responses and are subject to the "lock and key" notion of biological activity. Certain drugs in seemingly diverse chemical classes actually have a resemblance in spatial shape and charge distribution. It is not necessary that the whole molecule fit the receptor site to elicit a biological response but that only the "active portion" must fit the major portions of the receptor site. An application of this "lock and key" concept can be found in the development of PAM, an antidote for the toxic nerve poisons, which is effective against the alkyl phosphates - the most powerful chemical warfare agents (Ref. 5).

5.2.3 Absolute Sensitivity of Odor

Up until a few years ago, quantitative studies in olfaction were rare and no complete set of data was available with which the sensitivity of single olfactory receptors could be calculated.

The following six steps are involved in the calculation of the sensitivity of individual olfactory cells from the work of Stuver (Ref. 3).

1. The absolute threshold of the organ as a whole, that is, the minimum perceptible number of molecules, No entering the nose under optimal conditions
2. The fraction, f_1 , of the inhaled air that passes through the olfactory slit.
3. The fraction, f_2 , of odorous molecules left in the air when it reaches the olfactory slit. This factor, f_2 , is smaller than unity, because part of the molecules will be adsorbed on the mucous membrane that covers the nasal cavity.



4. The remaining molecules that pass through the olfactory slit will be effective only insofar as they hit the epithelium (fraction f_3). The number, N , of odorous molecules reaching the epithelium is then given by:

$$N = N_0 f_1 f_2 f_3 \quad (11)$$

5. From N and the number of sense cells in the epithelium it is possible to calculate the average number of molecules per sense cell. A simple statistical analysis gives the minimum number $\bar{n}$ of molecules to which a cell must respond. If one were to assume a high value of $\bar{n}$, stimulation at the actually observed threshold would never occur.
6. Finally, the steepness of the frequency-of-smelling curve gives information about the minimum number of molecules involved in the process.

A brief discussion of these six points follows:

The threshold depends on the time of presentation of the stimulus, concentration of the odorous substances, and rate of flow of the carrier gas (usually air). The rate of breathing is 250 cc/sec/ nostril. For short times of presentation of stimuli the threshold expressed in number of molecules is nearly constant but for longer stimuli it becomes proportional to the time of presentation which means that the concentration is the controlling factor.

The nasal cavity can be compared to a slit with a width of 1 to 2 mm. The area of the sensory epithelium or olfactory slit is 2.5 cm². For normal breathing it was found that 5 to 10% of the total flow passed through the olfactory slit, and thus provides an estimate of f_1 in Eq. (11).

It is known that odorous substances adhere strongly to the mucous membrane of the olfactory epithelium probably because the concentrations are so low that the amounts of odorous compound are not even large enough to leave a monomolecular layer on the surface. Assuming a complete loss of the odorous substance after collision with the wall, the amount a left after time t sec. is:

$$a = a_0 e^{-D \pi^2 t / d^2} \quad (12)$$



where d stands for width of slit (approximately 0.1 cm) and D for the coefficient of diffusion of the odorous substance (0.08) then

$$a = a_0 e^{-80t} \quad (13)$$

Where the rate of decrease is the same, whether the gases are at rest or flowing laminafly.

For the rate of breathing $V \text{ cm}^3 \text{ sec}^{-1}$, the time of transport from the entrance of the nose to the epithelium is about $2/V \text{ sec}$. Substituting in equation (12) gives:

$$a/a_0 = f_2 = e^{-(80 \times 2)/V} \quad (14)$$

since V (per nostril) is $250 \text{ cm}^3/\text{sec}$ then

$$f_2 = e^{-0.64} \approx 0.5 \quad (15)$$

substituting in equation (12), the fraction of molecules adsorbed is

$$f_3 = 1 - e^{-80t} \quad (16)$$

for normal breathing the result is

$$f_3 = 0.5 \quad (17)$$

substituting in equation (11) a value of 2% is obtained for the fraction of inhaled odorous material adsorbed in the olfactory slit. The other 98% is lost.

In order to calculate f_2 and f_3 the assumption was made that each odorous molecule hitting the mucous membrane adheres there for a fairly long time (of the order of a second or more.) However, this assumption does not appreciably affect the following conclusions.



The threshold for mercaptans is 10^9 molecules of which 2% or 2×10^7 are adsorbed effectively on approximately 4×10^7 sense cells which averages to 1/2 per cell. Assuming that n or more molecules are necessary to excite a cell, the probability is that at least one cell will be excited since one cell at least has to be excited in order to produce a sensation. If n is set too high no sensation would occur for threshold values that have actually been observed. The assumption has been made that the odorous molecules are evenly distributed over the sense cells. However, not all molecules hit effective sites and not all cells may be effective for the odor. Hence, this all leads to lower values of n . The value of n , derived under modified assumptions will be much lower than 8.

If one studies the steepness of the frequency-of-smell curve which relates the frequency of a sensation with the intensity of the stimulus, the steepness of the curve gives information about m , the number of molecules involved in producing a sensation. The data indicate m to be 40 which means that 40 cells may be necessary to produce an olfactory sensation assuming that each cell responds to a single molecule.

Thus, in summary, the threshold of one human olfactory cell is at most 8 molecules for appropriate odorous substances and that at least 40 molecules are necessary to produce a sensation.

5.3 Summary

Man is faced with many challenging problems in his attempts to design physical analogues of the biological receptors. The high sensitivity and short response time, characteristic of the more refined bioreceptors, have not been equaled by physical transducers. A few, or possibly even one molecule, by operating on specific chemoreceptors can cause an impulse to travel along a nerve fiber utilizing metabolic energy, and producing an electrical signal that can be detected by comparatively crude means. The site is then restored to its original state in a fraction of a second so that it is capable of firing more signals. This process is so efficient that biosensors are certainly worthy of detailed studies for potential applications. The gas chromatograph is the closest physical detection analogue to these chemoreceptors. It might be possible to (a) modify the separation column or absorber in a gas chromatograph in such a way that it is a chemical duplicate of the mucous membrane of the nose or even to incorporate the biosensor tissue itself into the instrument, and (b) to improve the detector



an recorder by introducing some computer technique to the output of the gas chromatograph so that an analytical tool would be available that would be capable of separating and identifying (after calibration) as many odors as the nose with the same high degree of sensitivity, a higher degree of objectivity, and an absence of fatigue.

Because of the high sensitivity of biosensors, these receptors would make excellent CW sensors if incorporated into physical transducer systems. The present document has shown potential valuable applications for investigations to determine the feasibility of using living tissue, organs or animals (insofar as the animal is supporting specific receptor tissues) as primary sensors. Along these channels, Allied Research has also been interested in CW agents not detected by the senses such as certain nerve gases, whose presence in very small quantities, are difficult to detect. Allied Research has proposed that, with the aid of a radioactive technique, methods for detecting nerve gases present in concentrations of less than 0.1 $\mu\text{g}/\text{m}^3$ of air can be devised. From such development, a method for a rapid continuous monitoring system suitable for field use is expected.

Some general avenues of approach for improving physical instruments should involve studies concerning the following biological structures and principles:

- a) investigations of semi-permeable membranes in order to find one that allows rapid diffusion,
- b) investigation of substrates having high adsorptive power for the molecules (stimulus) to be detected,
- c) an investigation of how adsorption "triggers" diffusion and,
- d) how diffusion generates voltage.

REFERENCES - SECTION 5

1. Comparative Animal Physiology, edited by C. L. Prosser, W. B. Saunders Company, Philadelphia (1950).
2. Bayliss, L. E., Principles of General Physiology, Vol. II, Longmans, Green and Co. Ltd. (1960).
3. De Vries, H., and M. Stuiver, "The Absolute Sensitivity of the Human Sense of Smell" in Sensory Communication, edited by W. A. Rosenblith, John Wiley and Sons, Inc., New York (1961) Chap. 9.
4. Geldard, F. A., The Human Senses, John Wiley and Sons, Inc. New York (1953).
5. Ferguson, L. N., and A. R. Lawrence, Journal Chem. Educ., Vol. 35, p. 436 (1958).
6. Moncrieff, R. W., The Chemical Senses, Leonard Hill Ltd., London (1946).
7. Pfaffmann, C., "The sense of taste" in Handbook of Physiology, Vol. I, Section I: Neurophysiology, edited by J. Field et al., Waverly Press, Inc., Baltimore (1959) Chap. XX.
8. Beidler, L. M., "Mechanisms of Gustatory and Olfactory Receptor Stimulation" in Sensory Communication, edited by W. A. Rosenblith, John Wiley and Sons, Inc., New York (1961) Chap. 8.
9. Beidler, L. M., American Scientist, Vol 49, p 421 (1961).
10. Hamor, G. H., Science, 134, 1416 (1961).
11. Gray, J. A. B., "Initiation of impulses at receptors" in Handbook of Physiology, Vol. I, Section I: Neurophysiology, edited by J. Field et al., Waverly Press, Inc., Baltimore (1959) Chap. IV.
12. Adey, W. R., "The sense of smell" in Handbook of Physiology, Vol. I, Section I: Neurophysiology edited by J. Field et al., Waverly Press, Inc. Baltimore (1959) Chap. XXI.



6. Thermoreceptors

Thermodetectors in the tissue of living things function as temperature sensors. They occur either as cold or hot detectors associated with a small sensitive area of the skin known as a spot. In general cold spots are far more numerous than warm spots. The distribution of warm and cold spots in human skin is given in Table 6-1. (Taken from Ref. 1).

TABLE 6-1
DISTRIBUTION OF WARM AND COLD SPOTS
IN HUMAN SKIN *

	Cold Spots	Warm Spots
Forehead	5.5-8	
Nose	8-13	1
Lips	16-19	
Other parts of face	8.5-9	1.7
Chest	9-10.5	0.3
Abdomen	8-12.5	
Back	7.8	
Upper arm	5-6.5	
Forearm	6-7.5	0.3-0.4
Back of hand	7.4	0.5
Palm of hand	1-5	0.4
Finger dorsal	7-9	1.7
Finger volar	2-4	1.6
Thigh	4.5-5.2	0.4
Calf	4.3-5.7	
Back of foot	5.6	
Sole of foot	3.4	

* Number per cm²

† After Strughold and Pors (Ref. 2)

‡ After Rein (Ref. 3)



"Concerning the thermal sensitivity of animals, our knowledge is very scarce and scattered. Until recently cold-blooded animals were not believed to possess any specific thermoceptive organs. Sand, using electrophysiological methods, discovered that the Lorenzian ampullae of Raja reacted to cooling. The Lorenzian ampullae of the elasmobranchs are situated laterally in the region of the head and consist of a group of small mucous cysts lying subcutaneously. They are supplied by afferent fibers from the facial nerve."

"The facial pits of the pit viper (Crotalidae), which originally were believed to function as mechanoreceptors specialized for the detection of air vibrations, were clearly shown by Noble and Schmidt through behavioral experiments to detect the body temperature of the snakes' prey. They proved that snakes with the other principal sense organs of the head nonfunctional can still strike correctly at moving objects and can discriminate between warm and cold ones as long as the pits are uncovered. The organ consists of a small pit about 3 mm in diameter covered by a membrane 15 μ thick. This thin membrane is the innervated sensory surface. Leading off from microelectrodes, steel needles with tip diameter of about 3 to 7 μ inserted into the membrane, Bullock and Cowles, Bullock and Diecke, and Bullock and Faulstick proved that the afferent nerve endings serve as infrared receptors. They are, so far as we know at present, the most densely distributed warm receptors and the most effective organ for infrared detection within the animal kingdom. In mammals cold sensitivity seems to be located particularly on the bare parts of the nose and on the tip of the tongue. More details are not available as yet." (Ref. 1)

In some parts of the skin the thermoreceptors have been found to be the Krause end bulb for cold and the Ruffini end organ for warmth. They are located at a depth from the surface of about 0.17 mm and 0.3 mm respectively. Whether the Krause end bulbs are the receptors for cold in other parts of the skin is still uncertain since conventional histological methods have failed to reveal any end bulbs of the Krause type in the skin underlying the cold spots.



The response time to a thermal stimulus comprises the time of propagation of the thermal wave through the skin, the latency of the thermodetector and the nerve transmission time. Typical figures for recordings made in the lingual nerve of a cat in response to a cold stimulus on the tongue are: total response time 0.015-0.07 sec, physiological latency 0.003 to 0.005 sec.

Threshold temperatures for onset of a warmth sensation averaged over 8 human-subject experiments, are given in Table 6-2. (Taken from Ref. 1).

TABLE 6-2
THRESHOLD TEMPERATURES FOR WARMTH
DURING HEATING OF BODY IN CLIMATE CHAMBER *

Rate of Increase degrees/sec	Forehead ^{°C}	Abdomen ^{°C}	Hand ^{°C}	Foot ^{°C}	Integrated Skin Temp ^{°C}
0.001	34.8 $\pm$ 0.3	34.5 $\pm$ 0.6	31.7 $\pm$ 1.1	31.5 $\pm$ 1.1	34.2 $\pm$ 0.6
0.002- 0.003	34.7 $\pm$ 0.5	34.8 $\pm$ 0.8	31.5 $\pm$ 1.9	31.5 $\pm$ 1.9	34.3 $\pm$ 0.94

* Average values from 8 experiments [From Marechaux and Schäffer (Ref 4)]

In conclusion it seems that there are not a great deal of data available concerning thermoreceptors in living organisms. From the limited data that have been surveyed it has not been possible to formulate any engineering applications. Generally speaking, adequate thermodetectors are available in our own technology. It is conceivable however that further study could lead to the generation of applications, particularly to reproduce the small size of bio thermodetectors.



REFERENCES - SECTION 6

1. Handbook of Physiology, Section I Neurophysiology, Vol 1, American Physiological Society.
2. Strughold, H. , Porz, R, Ztschr. Biol. vol 91, p 563 (1931).
3. Rein, H. , Ztschr. Biol. vol 82, pl82 (1925).
4. Marechaux, E. W. , Schafer, K. E. , Arch. ges. Physiol. vol 251, p 765 (1949).



7. Electric and Magnetic Receptors

The ability of certain fish to generate electrical discharges has been widely known for some time. More recently it has appeared that some of these electric fish are capable of locating objects in the water. The role of magnetism in animal behavior is neither as well understood nor universally accepted. The preliminary examination of the literature performed under the present study indicates an accelerated interest in both fields, and indications are that the coming year will bring a considerable amount of new data.

7.1 Electric Receptors

Among the animals inhabiting the earth, only fish, both marine and fresh water forms, are known to exhibit the ability to generate electrical pulses. This property is not restricted to a particular group of fish; it is found in both the elasmobranchs and teleosts. The generating organ itself is found in various locations on the body of the fish. It is found in the modified pectoral fin region of rays, and in the caudal region of bony fish.

The electric field serves two functions: in the case of Electrophorus, the South American electric eel, or Torpedo, the electric ray, its function is offense or defense. In the families Mormyridae and Gymnotidae the electrical discharges are apparently used for object location. For example, within the family Mormyridae, Gymnarchus niloticus continuously discharges low voltage electrical pulses into the surrounding medium for object detection. Mormyrids and gymnotids are most commonly found in those areas of Africa or South America where streams are generally very turbid during the dry seasons and increase in turbidity during the rainy season when soil erosion is at its maximum. Under these conditions normal ocular sensing (vision) is of little value, and because of this the eyes are very poorly developed. The evolutionary process has resulted in the formation of organs that produce electrical discharges; these organs arise from multi-nucleated muscle fibers.

These fishes that use their electrical discharges for object location also possess electrical sensors. Thus we may speak of an active object-locating system, the energy for which is furnished ultimately by metabolic processes within the fish. Indications are that this system possesses some interesting data processing capabilities. In the course of this investigation, no evidence has been found to date of the existence of any passive electrical detection systems.



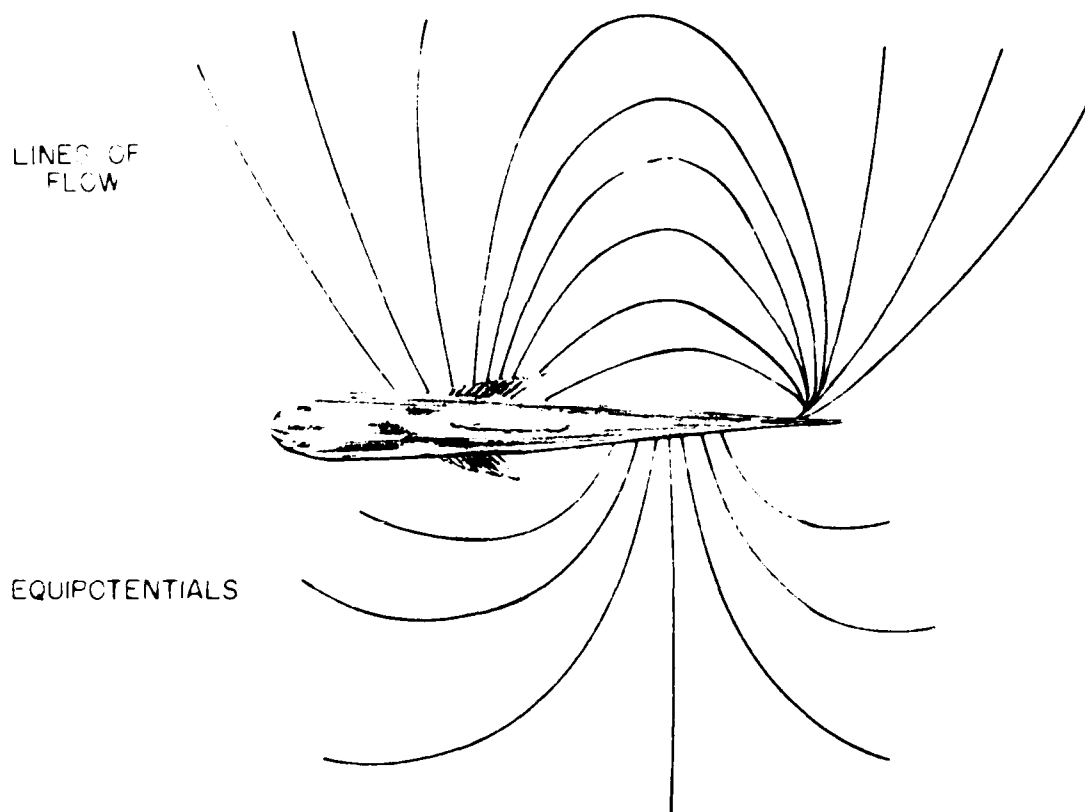
The object-locating system may be thought of in terms of three components - a transmitter, a receptor, and a data processing scheme. The transmitter apparently operates continuously, day and night, producing an endless pulse train of which the pulse duration and pulse frequency are different and characteristic for each species. The pulse frequencies range from slightly less than 50 cps to a few hundred cps. The frequency appears to be constant for a given fish and does not change when the presence of an object is detected.

The transmitting organs are found in the tail region lying parallel to the spinal column. In Gymnarchus there are four organs per side, while the other mormyrids exhibit two organs to a side. A transmitting organ is made up of many cells placed side by side anteroposteriorly. Each cell is separately innervated from the motor nerves of the central nervous system; it is surrounded by a gelatinous matrix which in turn is encapsulated by connective tissue. The electric impulse is believed to be triggered by a chemoresponse (acetylcholine) at the nerve ending of the individual organ cells.

The receptor organs have apparently been the subject of some controversy. It is now thought that the lateral line system, with its various ramifications over the head as far back as the posterior border of the opercular region, is an electrical receptor. Although no definite conclusions have been stated, certain glandular sense organs called mormyromast cells, found within the lateral line system, may be the true receptors. These organs may be able to sense any minute distortion of the electrical field established by the transmitter organ. The suggestion that mormyromast-type structures are electrical receptors was put forth by Lissmann (Ref. 1); these structures are situated in the skin and communicate with the surface by jelly-filled canals.

Consider the electric field generated by the transmitter electric organs. For a stationary fish in an infinitely large tank the lines of current flow and the equipotentials are similar to Figure 7-1, and correspond approximately to those of a dipole current source. When an object of a conductivity lower than that of the surrounding water is brought near, the pattern becomes like Figure 7-2A, while for an object of higher conductivity Figure 7-2F applies. The stimuli received by any electrical receptors on the surface of the fish will be modified (Ref. 1). Thus, it would appear that the presence of an object is detected through its influence on the electrical profile established over the skin of the fish by the transmitting electrical organs.

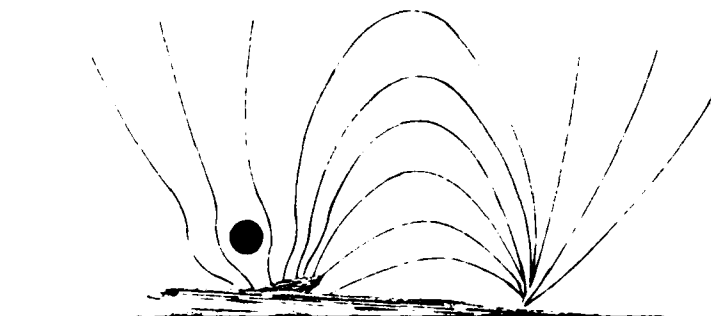
THE ELECTRIC FIELD AROUND A FISH



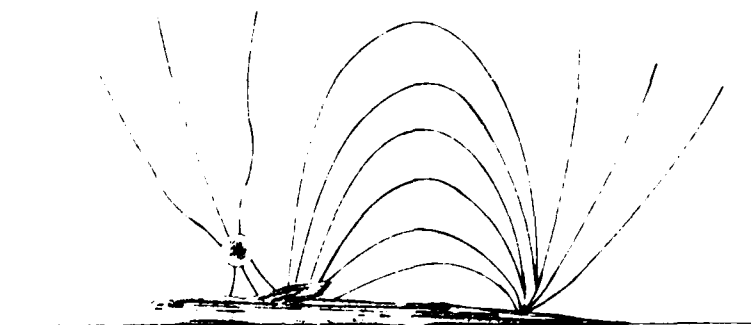


THE ELECTRIC FIELD IN THE PRESENCE OF AN OBJECT
(A) OF LOW CONDUCTIVITY, (B) OF HIGH CONDUCTIVITY

A



B





The third component of the detection system, the data processing scheme, is less well understood and may be the most interesting feature of the detection system. Lissmann and Machin (Ref. 1) performed experiments to find how the potential distribution around a fish would change in the presence of perturbing objects. A voltage was applied to an electrolytic tank at two points to simulate the dipole field of a fish. The potential in the tank was sampled by pickup electrodes arranged around these points in the shape of the body of the fish. With delicate and protracted measurements it was possible to locate a large insulating object a few centimeters away from the probe electrodes. However, with any apparatus of only moderate complexity it was impossible to detect changes in the potential distribution for small and remote objects which could easily be detected by Gymnarchus. It is thus suggested that a mere sampling of potential over the sensitized skin area is not sufficient to explain the performance of the fish in behavioral studies. One is led to suspect the existence of a more sophisticated sensing scheme.

The authors of Reference 1 undertook an analysis of electric field configurations for various arrangements of objects around the fish. In addition to calculating changes in potential caused by a perturbing object, they investigated the change of the second space derivative of potential. Their results led to the following conclusions which are quoted from Reference 1:

"(a) A characteristic disturbance occurs in both the potential and its second derivative around the point on the surface of the fish nearest to the object.

(b) The changes of potential are slowly varying and extend over almost the whole of the surface of the fish, while the changes of second derivative are much sharper and more closely confined to the part of the fish nearest to the object.

(c) In the second derivative mode the greatest sensitivity is obtained in the head-on position.

(d) In the head-on position the width of the peak of second derivative gives information about range: at short ranges a sharp peak flanked by troughs is produced, while at larger ranges the peak is broader. No such discrimination is available in the potential mode.



(e) For the second derivative mode the changes in stimulation of the receptors due to the presence of the object are of the same order of magnitude as the total stimulation they receive in the absence of the object. For the potential mode the changes are about one-fifth of the total stimulation.

(f) Movement of the tail of the fish produces relatively smaller changes of stimulation of the second derivative mode than for the potential mode.

It is clear that the second-derivative mode is capable of the greater location accuracy, and is less disturbed by tail movements. Furthermore, relatively greater changes of stimulation of the receptors are produced by operation in the second-derivative mode. " (Ref. 1).

The same paper describes an experimental program consisting of behavioral studies of a fish in response to externally applied electrical fields. It was concluded from the results of one of these experiments that Gymnarchus can detect objects by the disturbance of its own electric field in the water. The experiments described also indicate the limits of the sensitivity of the fish.

The authors (Ref. 1) conclude that the second-derivative mode appears to be the one most probably operating in Gymnarchus. The experimentally determined limits of detection are discussed in relation to the random noise in the receptor circuits; it is concluded that both spatial and temporal integration are likely to be employed. It is suggested that the detection system operates with a limited band width. This would be equivalent to an integration of the signal over a time τ given by $\tau = 1/(2 \pi \Delta f)$ so that the reduction of noise is achieved only at the expense of relatively long response time. It is further suggested that spatial integration is accomplished by averaging (possibly in the central nervous system) the response of many neighboring receptors. If n receptors are thus averaged, the noise can be reduced by a factor of $\sqrt{n}$.

The authors discussed two other mechanisms that could be employed to improve the signal-to-noise ratio for the receptor system as a whole. By swimming to and fro near the object the fish may 'scan' the area, giving a field pattern which sweeps over the receptors in a readily identifiable way. This gives an effective increase of integration time, since the information from the receptors may be collected over the time of one whole 'scan'. Furthermore,



the small relative change in stimulus would be much more readily detected since several comparisons could be made in a short time. The other mechanism involves inhibition of the receptors between transmitted pulses. For a one millisecond pulse with a repetition rate of 300/sec, an improvement of signal-to-noise ratio by a factor of about $\sqrt{3}$ would be obtained in this way. This 'blanking' would have to take place before the point at which temporal integration occurred; there seems to be no obvious mechanism for carrying this out. Since the signal-to-noise ratio is improved by a relatively small factor, it is unlikely that any very complicated mechanism would be involved to give 'blanking' of the receptors.

In another paper (Ref. 2) the same authors describe further experiments the results of which appear to confirm their postulation of the possible mode of action of the electric receptors.

The electric receptors are assumed to be the mormyromasts which must measure the amplitude of a signal consisting of one millisecond pulses at a repetition frequency of about 300 cps; these pulses are emitted by the electric (transmitter) organ. The information about the signal amplitude has to be coded and transmitted down a sensory nerve in the form of impulses with a maximum rate of the order of 500 cps. It was proposed in Reference 1 that this could only be done by 'smoothing' the incoming pulses and transmitting a sensory nerve signal characteristic of their mean value. An integration time constant of about one quarter second was suggested for the smoothing mechanism. This mechanism was referred to in Reference 2 as "pulse-frequency-modulation" and was compared therein with an alternative mechanism called "pulse-phase-modulation". The experiments described were designed to distinguish between the two mechanisms. The experiments consisted of behavioral studies of Gymnarchus in an electrolytic tank in which carbon plates were immersed for the purpose of applying electrical stimuli. The stimuli consisted of pulses of which the frequency, amplitude, and width were variable. The fish was trained to feed only after sensing an applied stimulus. Administration of an electrical stimulus above threshold brought about a marked increase in the swimming speed of the fish, a direct approach to the feeding trough, and entry from the side



farthest from the carbon plates. Thereupon the fish was fed. This was the "standard response". Application of a sub threshold stimulus caused no reaction and the fish continued to circle steadily. It was thus possible through a series of trials to establish the threshold of each of several pulses of different frequencies and pulse widths.

The threshold field for the standard response is plotted in Figure 7-3 as a function of pulse width for three frequencies. The straight lines of slope (-1) indicate the relation to be expected on the basis of the pulse-frequency-modulation hypothesis; their absolute position is chosen to give the best fit to all the results displayed. The horizontal line representing the threshold to d.c. agrees well with the pulse results, each sloping line cutting it at the appropriate point. Thus, d.c. can be regarded as one-millisecond pulses at 1,000 cps, or five millisecond pulses at 200 cps, etc. It is clear that the results are consistent with the pulse-frequency-modulation mechanism.

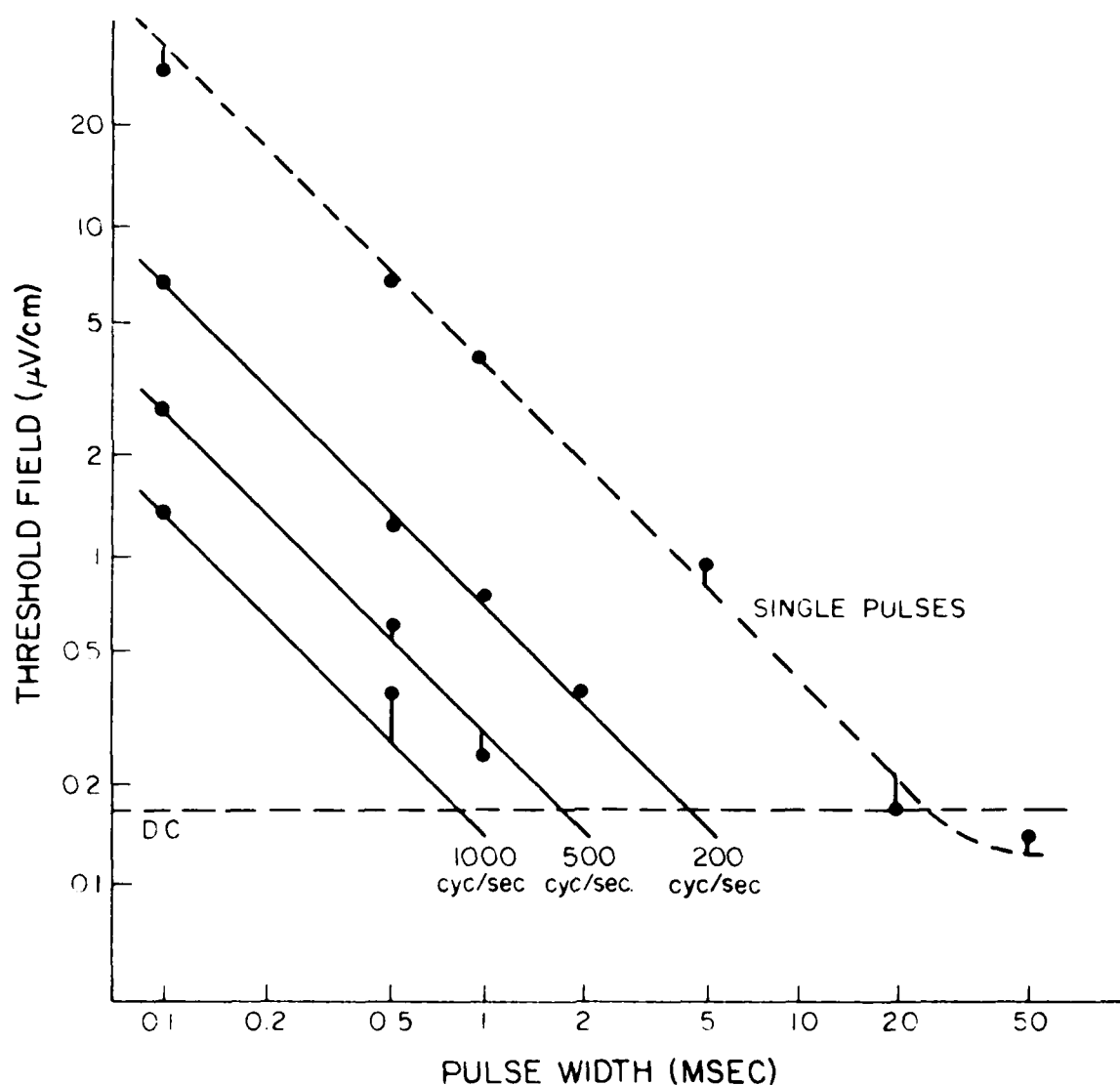
The dotted line in Fig. 7-3 gives the threshold for single pulses; the authors interpret it as further confirmation of the pulse-frequency-modulation hypothesis. From its position relative to those for repetitive stimulation, the integration time constant of the receptors is found to be 25 milliseconds.

The results for a pulse repetition frequency of 50 and 20 cps (not shown in the figure) indicate that when the pulse width or interval becomes comparable with the integration time constant, the threshold is no longer inversely proportional to pulse frequency. Indeed, the fish is more sensitive to 20 millisecond pulses at 20 cps than to d.c.

An explanation of this anomaly can be given on the pulse-frequency-modulation hypothesis by assuming that adaptation takes place before the transmission of information up the nerve to the CNS (central nervous system). When the stimulus is switched on, the receptors integrate with a time constant of 25 milliseconds, then send a signal to the CNS characteristic of the charge passed during the integration. If the stimulus is continued the signal gradually ceases. Thus the sensitivity of the fish to single 20 and 50 msec pulses is about the same as to d.c. (Fig. 7-3).



THRESHOLD FIELD AS A FUNCTION OF PULSE WIDTH AND FREQUENCY





If, however, the stimulus is 20 msec pulses at 20 cps, the receptors will send out a signal on the first pulse, reset during the interpulse interval, and send out a further signal for each subsequent pulse. Thus the effect at the CNS of pulses spaced sufficiently apart in time will be greater than that for a d.c. stimulus of the same voltage. The adaptation and resetting time is not likely to be less than the integration time, and the results from the experiments indicate that it is probably of the same order. The authors anticipated undertaking further research in this area (Ref. 2).

In reconciling their values of sensitivity with the observed integration time constants, the authors were forced to assume spatial integration over some 50,000 receptors. Mere averaging over such a large number would destroy some of the detail in the pattern of information about the receptor field. Accordingly, they suggest, it is necessary to assume that the spatial integration must take the form of a pattern recognition process whereby the pattern of information from the receptors is identified with one of a series of stored patterns, either innate or previously learned. Such a process, they state, is formally identical with spatial integration over a large number of receptors.

The suggestion that a pattern recognition process exists in these fish is perhaps the most interesting feature of the object-locating system. As pointed out in other sections of the present report, pattern recognition is one of the crucial areas in which we lack adequate understanding. Thus, although it is interesting that the individual electrical receptors can probably detect changes in potential gradient on the order of $0.15 \mu\text{V}/\text{cm}$, the potential opportunities for research in the area of pattern recognition may be more significant.

The electric receptors themselves provide an opportunity to study the general methods of sensor operation. In Reference 1 it was shown that the threshold of sensitivity of Gymnarchus corresponds to a change of current in a single receptor of $0.003 \mu\text{A}$. A current of this magnitude flowing for 25 msec (receptor integration time) corresponds to a movement of 1000 univalent electrons. The mechanism by which such a minute change in current can control the frequency of nerve impulses remains totally unexplained (Ref. 2).



Most of the studies of electric fish are very recent. It is clear that further advances in this area of knowledge will be of interest to the instrumentation engineer for its possibilities in introducing new mechanisms for low-level signal processing as well as new insight into the pattern recognition methods of nature.

7.2 Magnetic Sensors

Montgomery (Ref. 3) raises the interesting point that claims of magnetic field effects on living processes have historically been associated with charlatanism, and that to this day they are received in many quarters with skepticism. No one questions the effect of gravitational, electrical, or concentration-gradient fields on biological processes; yet magnetism, as fundamental a phenomenon as the others, gets short shrift. He expresses belief that the coming year will produce a body of new data that will support the notion of biological response to magnetic fields, and indicates his own inclination toward a positive view on the matter.

The material examined during this study (Refs. 4 through 6) consists of extensive behavioral studies of the snail under the influence of magnetic fields. Evidence is advanced to support the hypothesis that the orientation of snails normally includes a true response to the earth's magnetic field (Ref. 4). The evidence is based upon statistical analysis of many samples of the snails' motions.

It is tempting to contemplate the possibilities of applying sensitive biological magnetic detectors to such problems as submarine detection and navigation. From an engineering point of view, however, the research to date is of interest only in that it suggests that a useful natural phenomenon may exist. The fact that no specific sensor has been identified with magnetic behavior makes it clear that any engineering application of this biological process will have to be preceded by considerable and additional research.

REFERENCES - Section 7

1. Lissman, H. W., Machin, K. E., The Mechanism of Object Location in Gymnarchus niloticus and Similar Fish Journal of Experimental Biology, Vol. 35, pp. 451-486 (1958).
2. Machin, K. E., Lissman, H. W., The Mode of Operation of the Electric Receptors in Gymnarchus Niloticus, Journal of Experimental Biology, 37, 4 Vol. 37, No. 4, p. 801 (1960).
3. International Science and Technology, Prototype Issue, Letter from P. J. Montgomery of Michigan State University.
4. Brown, F.A. Jr., Brett, W. J., Bennett, M.F., Barnwell, F.H., Magnetic Response of an Organism and its Solar Relationships, Biological Bulletin, Vol. 118 No. 3, (1960).
5. Brown, F.A. Jr., Webb, H. M., Brett, W.J., Magnetic Response of an Organism and its Lunar Relationships, Biological Bulletin, Vol. 118, No. 3, (1960).
6. Brown, F.A. Jr., Bennett, M.F., Webb, H.M., A Magnetic Compass Response of an Organism Biological Bulletin Vol. 119, p. (1960).



8. Neuron Transmission

Nerve fibers in living organisms serve to convey information between the brain (as well as certain other nerve centers) and the various sensing organs and motor muscles throughout the body. The mechanism of conduction and transmission in nerve fibers is unlike that of electronic conduction employed in our technology.

The operation, even the very existence, of many of the interesting biosensors depends on the peculiar characteristics of the nerve fibers that connect them to the central nervous areas. Thus, for example, the eye is connected to the brain by one million nerve fibers, each of which carries a signal which is a response to a very small input stimulus, i.e. the arrival of 1-10 quanta at the eye. If the eye is to be duplicated in our technology, considerable attention will have to be devoted not only to the sensor itself, but also to the information linkages connecting it to the rest of the system for which the sensor is intended. Furthermore, the use of such an artificial eye as a pattern sensor - one of the most sought after facilities from the bio world - will surely require large scale logical interconnections between the sensory elements of the eye. This would be cumbersome using electronic conduction and discrete logical elements. With analogs of nerve fibers, logical circuits will be much easier to construct, since each segment of a nerve fiber is, in a sense, a logical element.

In this section the mechanism of neural conduction is described and a useful application in our own technology of this method of information conduction is discussed. Finally the general characteristics of man made nerve fibers are considered.

8.1 The Membrane Theory (Bernstein) of Nervous Condition (Taken from Ref. 1)

"Conduction, according to this theory, is a surface phenomenon. The nerve fiber is surrounded by a semi-permeable membrane or surface film which is polarized when the nerve is at rest. That is, the surface film separates a layer of cations on its outer side from a layer of anions on its inner side. A stimulus applied to the nerve increases the permeability of the membrane at



the point of stimulation with the result that a redistribution of ions and depolarization of the membrane occur. This point of the nerve becomes thereby relatively negative to the inactive (polarized) section of nerve immediately adjacent. A potential difference is set up and a current flows between the active and inactive sections. This secondary current in turn causes depolarization and activation of the adjoining region which, being now relatively negative to the next section, results in a current being again set up between these two which depolarizes the latter. Thus the currents set up between contiguous, inactive and active regions serve as a successive stimuli and the wave of depolarization spreads down the nerve; the disturbance set up by the original artificial stimulus is in this way propagated automatically (see fig. 8-1). The depolarized state persists for a short time after the passage of the impulse; during this time the nerve is refractory. The return of excitability is dependent upon the restoration of the polarized state. According to the membrane theory the nerve impulse is simply the propagated wave of depolarization."

"Perhaps the greatest support for the membrane theory is afforded by the experiments of Lillie who has prepared a metal model which behaves in a manner comparable to that of nerve. A film of oxide forms upon an iron wire placed in strong nitric acid. When such an oxide-coated wire is then immersed in a weak solution of acid, which would cause gradual solution of an untreated wire, no reaction results. The metal is in a so-called passive state, being comparable to a resting nerve. When, however, the wire is "stimulated," e.g., touched at one point with active iron or some other base metal, scratched with a piece of glass in order to break the protective oxide film, or an electric current applied to it, a reaction (electro-chemical reduction) is set up which, accompanied by effervescence and the formation of a dark-colored lower oxide, sweeps down the wire. If two parts of the wire be connected with a galvanometer a current flows through the instrument during the spread of the reaction. If the acid bathing the wire is of a certain concentration, the protective film reforms in the wake of the reaction. If a second "stimulus" is applied after the film has reformed, but not before, a repetition of the phenomenon occurs. The resemblance of the reaction of the iron

wire to the nerve impulse is very striking. Both are electrical in nature; the film of oxide is analogous to the surface film postulated for the nerve fiber. So long as the protective oxide is intact a potential difference exists between the wire on the inside and the surrounding acid. Discontinuity at any point in the film causes this region to become negative to other regions and a wave of depolarization is propagated to the end of the wire."

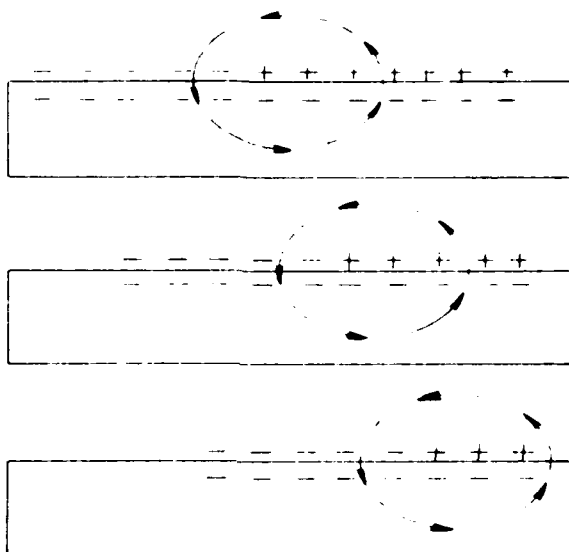


Fig. 8-1 Illustrating nervous conduction according to the membrane theory.

"In its behavior the wire model shows the following remarkable resemblances to that of nerve.

- (1) The rate of propagation of the reaction is of the same order as that of the nerve impulse.
- (2) The "stimulus" must be of a certain intensity (threshold). A stimulus greater than the threshold causes no greater effect ("all or none" principle).
- (3) The wire, as already pointed out, is irresponsive for a short time after the reaction has passed (absolute refractory period); while the film is reforming the reaction is set up with greater difficulty (relative refractory period).



(4) when an electric current is employed to activate the wire it must rise to its maximal intensity rapidly in order to elicit the reaction.

(5) The activating current must flow for a certain minimal length of time (chronaxie) in order to produce the effect."

"In the wire model a demonstrable temporary break in the continuity of the oxide deposit is produced. It is presumed that the excitation of a nerve is associated with a corresponding change (increase in permeability) in the surface film surrounding the nerve fiber. Though such a permeability change has been demonstrated in certain slowly conducting protoplasmic systems, the evidence for its occurrence in nerve is indirect. In the case of nerve and other protoplasmic systems the alteration in permeability is associated with chemical changes - the processes underlying the excitation of nerve and the transmission of the impulse are therefore, as in the case of the iron wire model, essentially electrochemical in nature" (Ref. 1).

8.2 Application of the Neuron Transmission System

The contrast between the neural mechanism of information conduction and the normal electronic method used in our technology is of great interest. Neuron transmission involves, essentially the propagation of an electrochemical anomaly. In some respects a nerve fiber may be thought of as the limiting form of a chain of flip-flops in which the number of these logical elements per unit length has become infinite. Individual nerve fibers are very small (e. g. the optic nerve trunk is only a few millimeters in diameter yet it contains about a million individual fibers). The materials out of which nerves are made do not include very good electronic conductors or insulators. The rate at which information is propagated along individual nerve fibers is relatively small, of the order of 10-100 bits per second. Thus, when large amounts of information have to be transmitted, as in



the case of the retina/brain link, recourse is made to the use of very large numbers of individual fibers. In our technology such large data rates are normally transmitted over a single high frequency cable.

The contrast in information transmitting techniques between our own and the bio-technology prompts the following two questions:

- a) Why are the techniques of our technology (which seem so satisfactory) not used in the bio-technology?
- b) Have the bio-techniques any useful application to our own technology?

The answer to the first question can only be but speculative. The fundamental differences between the two technologies, as discussed in Section 1.1, are probably relevant. The bio world seems to be devoid of good electronic conductors and insulators. High frequency techniques such as time division and frequency division multiplex are nowhere evident. Thus, it may simply be the case that the information transmitting techniques of the bio world are a natural consequence of its technological limitations.

However, the possibility remains that in some way the bio-techniques are inherently superior to our own, and this leads naturally to the second question posed above. To attempt an answer, it is necessary to postulate a specific application or criterion for making comparison between the transmission techniques. For this purpose the ability to transmit large amounts of information by means of a cable of very small cross sectional area will be considered here as an example. This is not in any way a general criterion of usefulness, but it is interesting, partly because of the inherent micro nature of the bio-technology, and also because very high data rates per unit cross sectional area of cable may well be required in computers of the future: the trend in computer technology is to smaller size and higher data rates and data volume.

There are many practical limitations involved in the design of information transmission systems. However, the only fundamental limitation considered here is thermal noise together with certain basic properties of transmission lines and cables.



Suppose that it is desired to transmit information between two points by means of a very thin wire. In order to accomplish this, a certain minimum signalling power is necessary so that the received signal may be at a desired level above thermal noise. Thermal noise power is given by

$$W_N = 4kTf$$

where k is Boltzmann's constant, T is the absolute temperature and f is the bandwidth over which the noise power is totalled.

Now with a very thin cable the problem of power dissipation in the cable, due to ohmic losses, is important. As the cable gets smaller its resistance gets higher and thus unless the load and generator impedances are made equally high, more and more of the signalling power will be dissipated in ohmic losses in the cable. However, making the terminal impedances high is not a profitable solution for the case of high frequency transmission. Except at low frequencies, **the cable characteristic impedance must be matched to its terminal impedances**, and if these are high then so also must be the cable impedance. A high cable characteristic impedance can be obtained by making the ratio of the diameter of the overall cable (e.g. the outer conductor of a coaxial cable) much larger than the diameter of the individual conductor (s) (e.g. the center conductor in a coaxial cable). Thus, if there is to be fixed value of overall cable cross sectional area, the impedance of the cable can be increased by using a conductor (s) with a cross sectional area much smaller than this available area, the rest of the space being filled with air or other dielectric. Thus, any attempt to make a high impedance cable results in an increased value of ohmic resistance which in turn requires higher terminal impedances. Beyond a certain point this solution is self defeating. What is required, of course, is a minimum value of the ratio:

$$\frac{\text{ohmic resistance of cable conductor}}{\text{characteristic impedance of cable}}$$



for a given overall cable size. This minimum value exists, and in the case of a coaxial cable, occurs when the characteristic impedance is $77 \sqrt{\epsilon/\epsilon_0}$ ohms where ϵ is the permittivity of the dielectric and ϵ_0 is the free space permittivity.

Thus, with high frequency transmission, most of the signalling power will be dissipated in ohmic losses in the cable as soon as the ohmic resistance of the cable is higher than 100-1000 ohms.

The thinner the cable the more power will have to be transmitted – in order that the **received signal may be above the noise. A higher power requires a higher voltage.** However, as the cable is made smaller its power handling and voltage handling capacity will decrease. Thus, there is a physical limitation to how small a signalling cable can be made. It has been shown how this limitation involves inherent thermal noise and the thermal and electrical properties of the material out of which the cable is made.

With low frequency signalling the cable characteristic impedance does not have to be matched to its terminations. However, general limitations can be discovered to limit the degree to which the size of the cable can be reduced. Thus, as the cable is made smaller, its ohmic resistance will increase, the resistance of the terminations must be increased, the voltage for a given transmitted power level will increase – and at the same time, the breakdown voltage for the cable will decrease because of its small size. Also, the self capacity and inductance of the cable will, in relation to the impedance of the terminations, restrict the frequency bandwidth – which is another way of saying that at high frequencies the cable and its terminations must be matched.

These general considerations can be made clearer by considering a specific example. Suppose the cable is 1 meter long with a diameter of 10^{-8} meters. The D.C. resistance of this wire will be

$$R = \frac{\sigma l}{A} = \frac{1.7 \times 10^{-8} \times 1}{\pi/4 \ 10^{-16}} \text{ ohms}$$

$$\sim 2 \times 10^8 \text{ ohms}$$

If the cable is coaxial, or in the form of a strip transmission line, its characteristic impedance will be of the order of 100 ohms (for minimum loss). The



attenuation due to Ohmic losses (as a voltage ratio) of such a cable is given by:

$$e^{\frac{2 \times 10^8 x}{10^2}} = e^{2 \times 10^6 x}$$

where x is the length of the cable. For lengths greater than 10^{-7} meters, the attenuation is very great. In order to receive over a length of one meter (without the use of repeaters) a signal power equal to the thermal noise power, the transmitted power must be

$$e^{4 \times 10^6} \times W_N = e^{4 \times 10^6} 1.6 \times 10^{-20} \text{ watts}$$

(where W_N has been evaluated at room temperature). Thus, for any practical bandwidth (f) the transmitted power level must be impossibly high. It is clear, of course, that this type of cable can only be used with repeaters every 10^{-7} - 10^{-6} meters. (It can be shown that the total repeater power is minimized when the attenuation between repeaters is e times).

Consider now this same cable used for low frequency signalling between impedances of approximately 2×10^8 ohms. The signalling voltage is given by

$$V^2 \sim 1.6 \times 10^{-20} f 10^8$$

$$\text{i.e. } V \sim 10^{-6} \sqrt{f} \text{ volts}$$

The potential gradient in the cable is then

$$V/10^{-8} \sim 10^2 \sqrt{f} \text{ volts/meter}$$

Thus, voltage breakdown will be a limitation only at very large bandwidths (f). As shown below other limitations are important for much lower values of f .

The capacity of a coaxial cable or any two wire line is given, approximately, by

$$C \sim \frac{1}{20 \log_e \frac{b}{a}} \frac{\epsilon}{\epsilon_0} \times 10^{-9} \text{ farads/meter}$$



where b is the overall cable diameter and a is the conductor size. A low capacity requires a high value of b/a , which for a given value of b leads to a **small value for** a . This is the same situation as discussed earlier for characteristic impedance, and there again is an optimum value of b/a ($= 3.6$). Thus,

$$C \sim \frac{1}{20} 10^{-9} \text{ farads/meter}$$

For a one meter length of cable

$$C \sim 50 \times 10^{-12} \text{ farads}$$

The frequency at which the DC resistance of the cable (2×10^8 ohms) equals the reactive impedance of the self capacity is given by

$$\omega = \frac{1}{5 \times 10^{-11} \times 2 \times 10^8} = 10^2$$

i.e.

$$f \sim 16 \text{ cps}$$

At frequencies much above this value, the self capacity of the cable is important, and will give rise to (linear) distortion and attenuation.

It is concluded that satisfactory transmission over such very thin cable, of all but very low information rates, is possible only when larger numbers of repeaters are used. Such a cable, with repeaters, would, of course, be a very close analog of a nerve fiber. Thus, it is apparent that the neuron technique of information transmission may well have an application in our technology. The requirement for this type of **wire and associated fabrication techniques probably do** not exist at the present time, but are very likely to within a decade.

As mentioned previously a capability in neuron techniques will be required if ever close analogs of bio sensors such as the eye are to be built. The same problems will exist as discussed above, i.e. the propagation of large amounts of information over very thin cables.



8.3 The Neuristor

In the preceeding part of this section it has been shown that a useful application may soon exist for a man made version of a nerve fiber. Such a device has already been postulated (Ref. 2), and named the neuristor.

"A neuristor may be visualized as a one-dimensional channel along which signals may flow. A signal propagates along the channel in the form of a discharge, followed by a "refractory" period, during which time a second discharge cannot be supported. The mode of discharge propagation is somewhat analogous to that which occurs along a (pyrotechnic) "fuse," except that in the neuristor the channel exhibits perfect "healing". The refractory period can be thought of as the time of healing."

"A discharge signal has the following properties: 1) attenuation free propagation, 2) uniform velocity of propagation, 3) a refractory period."

"These characteristics are somewhat similar to the gross properties of transmission of discharge pulses by neurons in the nervous system - hence the name neuristor. A neuristor may be thought of as an electronic counterpart of the ionic neuron (in particular, the portion of the neuron that carries the propagating discharge)."

"In a sense, a neuristor may be considered to be a distributed version of a chain of suitably interconnected lumped-circuit monostable circuits. Devices of this type may be realized in many physical forms and may be based on many physical phenomena."

"The basic requirements for the realization of such a device are: 1) a distributed energy source, 2) a distributed energy storage, 3) a distributed active device" (Ref. 22).

The following conditions are considered to define, in general terms, a one dimensional neuristor.

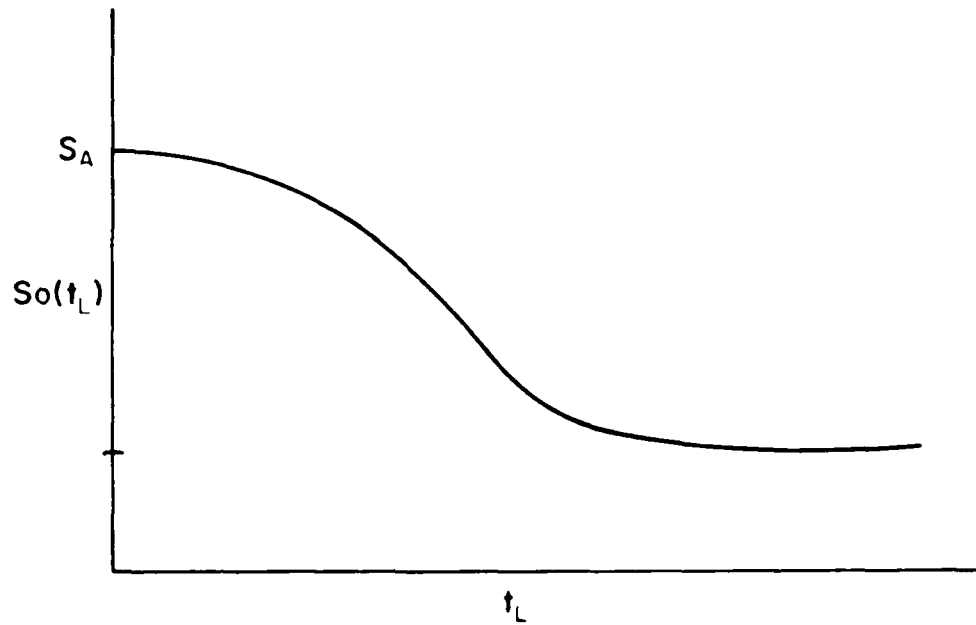


1. The neuristor is thought of as a uniform thin wire, the properties of which (e. g. temperature, voltage) vary only along its axis.
2. A quality, or state, of the wire is recognized and denoted numerically by S . (e. g., S could be temperature in degrees centigrade or volts).
3. Associated with this state is a current or flux denoted by f . (e. g., f could be calories per second or amps).
4. Laws of propagation in the wire are defined. For example,
 - (i) $f = -k_1 \frac{dS}{dx}$ (e. g. Ohm's Law)
 - (ii) $\frac{dS}{dt} = -k_2 \frac{df}{dx}$
5. When $S > S_0$ the state of the wire changes immediately to S_A . This will be an active (rather than passive) effect and will involve an influx of energy.
6. The level, S_0 , at which this unstable change occurs is a function of the time (t_L) since it last occurred. (See **Fig. 8.2**). **Thus, immediately after an unstable change in the state of the wire, the conditions necessary for the maintenance of S at S_A are destroyed.**

It is to be noted that the energy contained in a pulse transmitted down a neuristor (or a nerve fiber) does not come from the input to the neuristor cable, as is the case for a pulse on an ordinary coaxial cable. It comes rather from the active device or medium associated with the neuristor. At any time, the pulse itself will rapidly die out due to dissipation. It is however, continuously re-generated by the active mechanism. Thus, during the conduction of signals, power is being supplied continuously by the active medium and in this sense the medium resembles the DC power supply of an amplifier. In the numerical example quoted above, the power dissipation in the very thin copper cable was shown to be according to the law



POSTULATED NEURISTOR RECOVERY CURVE





$$P_{out} = P_{in} e^{-2 \times 10^6 x}$$

Thus, the space rate of loss of power is

$$\frac{d P_{out}}{dx} = - P_{in} \times 2 \times 10^6 e^{-2 \times 10^6 x}$$

Thus, to maintain a constant power level equal, say the thermal noise level, W_N , the repeaters must generate

$$W_N \times 2 \times 10^6 \text{ watts per meter length of cable}$$

$$1.6 \times 10^{-20} \times f \times 2 \times 10^6 \text{ (at room temperature)}$$

$$3.2 \times 10^{-14} f \text{ watts/per meter of cable.}$$

The volume of the cable is 10^{-16} cubic meters per meter length, thus, the repeater power may be expressed as

$$3.2 \times 10^2 \times f \text{ watts per cubic meter of cable.}$$

It is interesting to compare this figure with the quoted value of heat generation in a nerve. The resting sciatic nerve of the frog in oxygen generates about 70×10^{-6} calories per gram of nerve per second (Ref. 1). During stimulation in oxygen, at the rate of 280 shocks per second, an increase in heat production occurs of 40×10^{-6} calories per gram of nerve per second. Taking the density of nerve fiber as 1 gram/cc, stimulation at 280 shocks per second requires 4×10^{-5} calories per cc i.e., $4.19 \times 4 \times 10^{-5}$ watts per cubic centimeter i.e., 16.8×10 watts per cubic meter. Dividing this by 280 to introduce the bandwidth f , this figure becomes $6 \times 10^{-1} \times f$ watts/cubic meter - which is really quite close to the figure $3.2 \times 10^2 \times f$ watts/cubic meter for the thin copper wire.

8.4 Practical Realization of a Neuristor

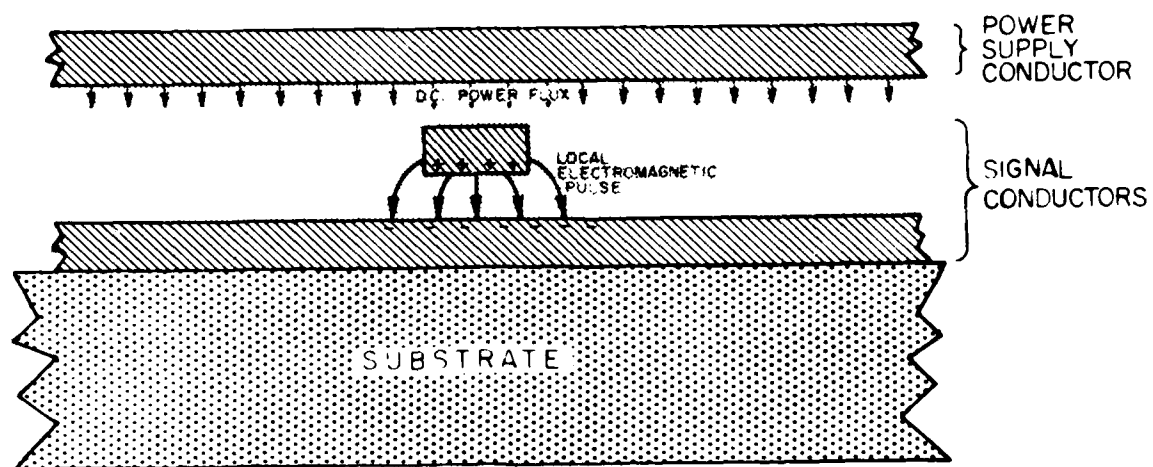
What is in effect a neuristor has already been described, namely the iron wire in nitric acid.



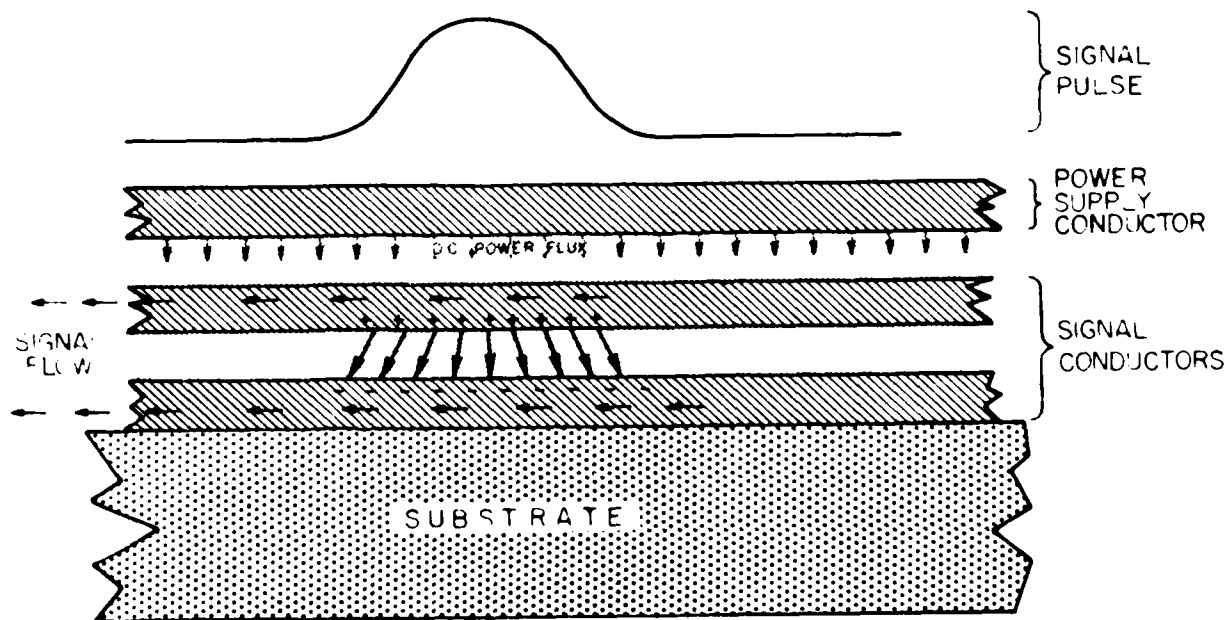
For applications postulated in this report the neuristor "wires" should be very thin and capable of being packed very closely together. One of the most promising methods of construction would be by the deposition of thin films, in strip form, as illustrated in figure 8.3. It is postulated that the necessary active and non linear characteristic of the neuristor (properties 5 and 6 of the list given in Section 8.3 can be attained by some complex of semiconductor materials filling the space between the two signal conductors and the power supply conductor.



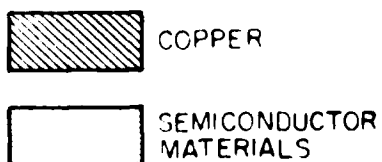
NEURISTOR SECTIONS



CROSS SECTION VIEW



SIDE SECTION VIEW



0 10⁻⁷m



REFERENCES - SECTION 8

1. Best, C. H., Taylor, N.B. Physiological Basis of Medical Practice, 5th Edition .
2. Crane, H.D. The Neuristor, IRE Transactions on Electronic Computers, September, 1960, p. 370.



9. Summary

The study reported here represents a substantial beginning in what should be a continuing effort to organize the pertinent data from the biological sciences into a body of knowledge that is meaningful in the field of instrumentation. The exploitation of biomechanisms in this manner has a potential of yielding superior classes of instruments that may be applied to the rapidly broadening spectrum of parameters to which mankind will so shortly be exposed.

It is not overdramatizing to say that a new era of instrumentation is at hand. At the present moment there are at least three frontiers beyond which there are countless new instrumentation requirements not yet conceived. The most obvious of these is the exploration of space, both by manned expeditions and automated instrument probes. The new environment, the number and complexity of desired measurements, and the necessity for long life with maximum economy of power and payload, all present us with a challenge to utilize all techniques at our disposal.

A second example, perhaps a little less demanding of our skill in miniaturization, is that of underseas exploration. It has recently become widely evident that a world as strange and unknown as outer space resides in the depths of our oceans. New forms of life, vast sources of food and energy and other vital resources abound here. It is likely that this domain will provide both the new variables to be measured, and in some cases the new means of measurement.

The third example exists physically right in our midst. Man has barely discovered a frontier in the realm of data handling and processing. Although several order-of-magnitude advances have been made since the first electronic digital computers of little more than a decade ago, they are child's play when put up against our rapidly unfolding awareness of things that ought to be done. The early computers so impressed the world with their facility that they acquired the name "giant brain." It quickly became apparent that this name represented their most dire shortcoming -- they were indeed giant. Yet today, scaled down many decades in physical size, they are barely able to be applied to the crudest attempts at imitating the thought processes or the pattern-recognition characteristics of animals.



To these examples, and other frontiers that they represent, the application of biomechanisms or their models seems inevitable. The immediate methods of application are, unfortunately, not always apparent. The purpose of the present study has been to establish some preliminary guidelines for an interdisciplinary endeavor of this sort, and to accumulate a body of data with which to begin.



10. Recommendations

The preliminary nature of this study has been emphasized at several points in this report. It is readily apparent that additional work, beyond that already done, will be needed if the potential value of this program is to be realized. A substantial investment has already been made in developing an approach philosophy and establishing suitable communication between the several disciplines involved. This investment provides the basis for a continued and accelerated analysis of the total spectrum of biological sensors, covering the broad range of sensed parameters as well as the individual variations in sensors from phylum to phylum. Such a continued effort would provide the opportunity to study and evaluate more fully some of the mechanical, electronic, and chemical variations that could be scanned only cursorily in the work to date. It would also permit a more specifically detailed investigation of other areas.

In order, then, that the total potential of this interdisciplinary endeavor may be fully exploited, it is recommended that the study be continued and extended, particularly to include:

1. Further development of predesign and design criteria.
2. Further work on the bibliography.
 - a. keeping it current and supplementing weaker areas.
 - b. enlarging the scope to include reports on government sponsored work, and foreign literature, especially Russian.
3. A more careful look at less familiar categories of the sensed parameters as well as phylum differences.
4. A look at biosensors in the context of their relations to other bioprocesses, i.e., system organization.
5. A more detailed investigation of design possibilities for specific application to manned space flight.