Massachusetts Institute of Technology
Artificial Intelligence Laboratory
and
Center for Biological Information Processing

A.I. Memo 846
C.B.I.P. Memo 014

August, 1985

THE COMPUTATIONAL APPROACH TO VISION AND MOTOR CONTROL

Ellen C. Hildreth and John M. Hollerbach

Abstract. Over the past decade, it has become increasingly clear that to understand the brain, we must study not only its biochemical and biophysical mechanisms and its outward perceptual and physical behavior. We also must study the brain at a theoretical level that investigates the *computations* that are necessary to perform its functions. The control of movements such as reaching, grasping and manipulating objects requires complex mechanisms that elaborate information from many sensors and control the forces generated by a large number of muscles. The act of seeing, which intuitively seems so simple and effortless, requires information processing whose complexity we are just beginning to grasp. A *computational approach* to the study of vision and motor control has evolved within the field of Artificial Intelligence, which inquires directly into the nature of the information processing that is required to perform complex visual and motor tasks. This paper discusses a particular view of the computational approach and its relevance to experimental neuroscience.

Acknowledgments. Support for this research was provided for JMH by NIH Research Grant AM26710 and for ECH by Air Force Office of Scientific Research contract F49620-83-C-0135. This chapter also describes research supported by National Science Foundation Grant 79-23110MCS and a grant from the Sloan Foundation. This document was produced using the facilities of the Artificial Intelligence Laboratory, supported in part by DARPA contract N00014-80-C-0505.

Contents

1 Introduction

1.1 The Founding Principles of Artificial Intelligence

The computational approach to vision and motor control is an outgrowth of the field of Artificial Intelligence, from which the basic tenets are derived (Minsky, 1968). Research in Artificial Intelligence has two main goals: to develop computer systems that exhibit intelligent behavior and to understand the nature of intelligence itself. The field is founded upon two basic principles. The first is to separate the tasks performed by a complex information processing system from the hardware that carries them out. The second is to analyze natural intelligent systems through the synthesis of artificial systems that perform the same tasks.

The birth of computers led to a distinction between a process as specified by software and the machinery or hardware that executes the process. Computers over the decades have been built from a variety of components, including cams, relays, analog circuitry, transistors and microchips, yet all are capable of performing essentially the same computations. The thought naturally arises that neurons can be viewed as another form of computational machinery and that an intelligent process need not be limited to biological nervous systems, but could in principle be implemented in computers as well. If intelligent processes can be separated from hardware, then intelligence can be considered an abstract entity, subject to its own rules, laws and structure, and can be studied in its own right. These laws are independent of the underlying computational machinery and reflect fundamental properties of particular information processing tasks.

The birth of computers also presented the opportunity to duplicate an intelligent process in a machine; this capability has become a fundamental tool of Artificial Intelligence. The synthesis of intelligent processes leads to insights that otherwise cannot readily be obtained. In laying out a computer program to perform a particular task, all details must be resolved and hidden assumptions made explicit. Even if the program never performs the task successfully, the act of carefully specifying the steps of a process forces a rigorous analysis of the problem. If a theory of how to solve a problem can be embodied in a computer program, then the theory can be tested by demonstrating whether the program can solve the problem. The implementation of proposed methods for solving problems in vision and motor control often exposes important gaps in our understanding of these problems and sometimes reveals essential features of their solution that can radically transform our thinking. Thus, machine synthesis adds a critical hypothesis–and–test loop to the study of intelligent processes and can sometimes lead to serendipitous discovery.

The attempt to duplicate intelligent processes such as vision and manipulation has shown how surprisingly difficult these problems are to solve. Biological systems, which currently provide the only working examples of intelligent behavior, supply useful hints about solutions to these problems. Accordingly Artificial Intelligence has always included as goals an understanding of human intelligence, as well as the development of intelligent processes that run on a computer.

Artificial Intelligence arose in part from the feeling that the methodologies of experimental psychology and physiology by themselves were limited in their ability to yield deep insights into the functioning of the human brain. The premise was that one cannot determine how a complex system works simply by extrapolating from the properties of its elementary components. It is necessary to have a theory of what the complex system is trying to do and how it could be doing it, before the elementary components can be identified and fit together.

The example of bird flight illustrates the difficulty of understanding a complex system by only making observations on its behavior. Suppose one is interested in how birds fly. Pulling the feathers off a bird causes the bird not to fly; this observation might lead to the conclusion that the secret of flight lies in feathers. Research might then continue through study of the properties of feathers. In reality, it is argued that bird flight could not have been understood until the development of aeronautics. Through the attempt to build a flying machine, a set of physical principles was derived that then shed light on how birds fly and what role feathers play. In essence, there are many ways of realizing flight and feathers may just be an implementation detail for birds.

The principle of analyzing a complex system by duplicating its abilities emphasizes that ideas must work in principle. Hypotheses are sometimes put forth for vision and motor control that cannot work because they are too vague or are ineffective procedures. It is often the case that before we can understand how a biological system solves an information processing problem, we must understand in sufficient detail at least one way that the problem can be solved, whether or not it is a solution for the biological system. In effect, we may need to be engineers before we can be scientists.

The above suggests that it may be desirable to have available a set of *competence theories* before attempting to develop a *performance theory.* This distinction is borrowed from Chomsky (1965), who defined a competence theory in natural language as a grammar that generates correct sentences. A performance theory is a competence theory that generates sentences the way humans do. One is ultimately interested in performance theories for biological systems, but it may not be possible to develop a performance theory directly without first having available a number of competence theories. Compe-

tence theories provide bases for understanding a problem, through development of a set of concepts, principles, and procedures that can be drawn upon in particularizing to a performance theory. The criticism sometimes made, that computer implementations of intelligent processes are implausible biological models might be beside the point, insofar as such implementations may lead to competence theories that teach us more about the problem.

Developments in many fields contribute towards understanding human intelligence, and fields like mechanical engineering have emphasized the approach of learning by duplicating. What is the unique contribution of Artificial Intelligence that differs from mathematics, physics, psychology, control theory and engineering? Why did research in vision, manipulation and robotics arise in Artificial Intelligence laboratories? Computer Science and Artificial Intelligence have contributed a rich set of computational metaphors that already are entrenched firmly in daily language. There is no longer any question of whether metaphors such as representations and algorithms are relevant to understanding cognition or complex information processes. It is sometimes difficult to distinguish a cognitive scientist or linguist from a researcher in Artificial Intelligence. In addition, while the knowledge of Newton and Euclid is old, looking at geometry and physics in a computational framework is new. Vision and motor control studies place new demands on these areas and considerations of algorithms and computational complexity often force a reanalysis of how to formulate a problem.

The ultimate strength of Artificial Intelligence may not lie in its particular methodologies of separating algorithm from hardware and synthesizing artificial systems, however, but in the freedom to approach information processing problems without preconceptions. Artificial Intelligence is a young field and has not yet developed a rigid set of formalisms or approaches that predispose one towards viewing a problem in a single way. Artificial Intelligence research borrows from many fields, and this flexibility is essential for progress in inherently multidisciplinary undertakings such as the study of vision and motor control. Artificial Intelligence does not substitute for the necessary and important research in the fields of experimental psychology and neuroscience. Rather it complements these fields, and through a symbiotic interaction with them, can facilitate progress in the study of biological systems.

1.2 The Computational Approach to Neuroscience

The computational approach to neuroscience is essentially a top–down approach, emphasizing the importance of understanding the detailed nature of the problems posed by particular information processing tasks. There are at least three specific contribu-

tions that this approach can make. First, by elucidating the problems that need to be solved in vision and motor control, computational studies can aid the initial exploration of the function of neurons in the visual and motor pathways. Second, by elucidating the possible methods by which visual and motor tasks can be accomplished, computational studies can refine models of how neurons function and by what mechanisms. Third, computational studies can provide a powerful predictive tool. If a model for the function of a class of neurons is specified in sufficient detail to be implemented on a computer, then the behavior of the model can be compared directly with physiological data in a rigorous manner.

The computational approach to the study of biological systems was elegantly cast by David Marr into a framework of *natural computation* (Marr, 1982; Marr and Poggio, 1977), derived from the founding principles of Artificial Intelligence. Marr was attracted to the field of Artificial Intelligence after experiencing the limitations of traditional approaches to brain research in his early work on the cerebellar cortex. Marr had hypothesized a model for cerebellar function as implementing a simple form of associative memory (Marr, 1969). Yet he abandoned this line of research after realizing that this simple memory function was useful in a variety of computations, but shed no light on how complex motor behavior can actually be achieved.

In his later work in computational vision, Marr elucidated three distinct levels of analysis that are necessary for understanding an information processing problem:

1. A *computational theory* clarifies what problem is being solved and why, and investigates the natural constraints that the physical world imposes on the solution to the problem.

2. An *algorithm* is a detailed step–by–step procedure that represents one method for yielding the solution indicated by the theory.

3. An *implementation* is a physical realization of the algorithm by some mechanism or hardware.

These levels could be construed as a prescription for conducting research on complex problems: one first formulates a theory, then proposes an algorithm, and lastly designs an implementation:

$$\text{theory} \Rightarrow \text{algorithm} \Rightarrow \text{mechanism.}$$

In reality, problems are not solved in this rigid manner because constraints exist at all levels. Relevant experimental data, known properties of the biological machinery,

and the biological feasibility of algorithms must all be taken into account. Instead, the formula is best considered as a prescription for clear thinking about complex information processing systems. In essence, the computational approach regards an understanding of a problem in vision or motor control to be complete only when the problem can be explained at all three levels. When pursuing a particular line of research, it is essential to know which level is being addressed.

Thus the computational approach to neuroscience emphasizes the use of all sources of constraints: external constraints imposed by the task, constraints imposed by the biological machinery such as limbs and muscles, and constraints imposed by neuronal computing abilities. For example, the slowness of the proprioceptive feedback loops in biological motor control makes inapplicable many engineering control theories that rely on near–instantaneous feedback, although other aspects of modern and classical control theory are quite useful in analyzing biological motor control. Properties of biological systems may not only proscribe but also prescribe theories. For example, springlike properties of muscle can suggest mechanisms of trajectory control. Synaptic properties suggest the basic computational elements out of which algorithms are built (Koch and Poggio, 1984).

Finally, some problems may lack an available theory or may be so complex that we must look to biology for clues. It may be that these problems cannot be understood independently of the biological solution. Ultimately, a deep understanding of vision and motor control at the three levels of theory, algorithm and implementation, requires a strong bridge between experimental and theoretical studies of these problems. The flow of information is therefore in both directions:

$$\text{theory} \quad \Rightarrow \quad \text{algorithm} \quad \Leftarrow \quad \text{mechanism.}$$

The phrase *computational approach* has also been applied to certain neural modeling approaches that study how neural networks can operate and how these operations can be extrapolated to explain higher brain functions. Examples of this approach include the work on perceptrons (Minsky and Papert, 1969) and parallel "connectionist" networks (Ballard, 1985), as well as Marr's original work on the cerebellum. The word *computation* in this case refers to the detailed working of the processing hardware rather than to the algorithm that is executed by the hardware; hence the two approaches differ considerably. Of course an explanation of how the neural machinery operates is necessary for understanding biological intelligence and eventually algorithms must be couched in terms of elementary neuronal operations. Koch and Poggio (1984), for example, have proposed such operations from biophysical studies of dendritic trees, and have suggested

how some low-level vision algorithms could be implemented by networks that execute these operations. The computational approach described in this chapter stresses the need to consider both the problems that must be solved by the biological system and the properties of the neural hardware that implements the necessary computations.

The usefulness of detailed neural modeling for understanding the nature of the computations that are carried out in biological hardware depends in part on the specificity of the computation performed by the neural circuitry. Suppose a given neural network were capable of performing a general purpose computation, analogous to modern computers and also, as proposed by Marr, functioned for the cerebellum as mentioned above. Then it might be impossible to deduce what computations are taking place at a particular time, simply by recording the output signals of individual neurons. In the same way, it would be impossible to determine what computations are taking place in a modern computer, simply by recording voltages in the electronic circuitry. The behavior of the circuits is being analyzed at a level that is inappropriate for understanding the computations being performed. Suppose on the other hand that the neural network is closely tied to a particular computation.[1] Then the pattern of connections between individual neurons in the network and the electrical signals they carry might provide useful information about the computation. Even then, it might still be difficult to infer how the neural code represents information that is useful in the task being performed and how the computation is distributed over single cells, neuronal clusters, or even patches of dendritic trees.

1.3 Relation to Other Areas of Artificial Intelligence

Vision, manipulation, and robotics have been among the most successful areas for exploration by Artificial Intelligence, along with natural language. These areas possess an advantage over more cognitive domains such as learning, knowledge representation and reasoning, in that they represent the results of processing by neural mechanisms that lie close to the periphery. As a result, external constraints of geometry and physics can be brought to bear, making hypotheses more suitable for implementation and testing. In cognitive areas, hypotheses are more difficult to detail and to evaluate and there are fewer constraints on hypothesis formation to guide this research toward clear conclusions. Vision and motor control were chosen as a focus for this chapter, in part because of the relative success of research in these areas and because their study has established a strong bridge between Artificial Intelligence and the experimental neurosciences.

[1]Expressed in modern computer terms, suppose that a particular computation is *compiled* into special-purpose electronic hardware.

On the other hand, it is by no means true that vision and manipulation offer simpler problems than those posed by higher cognitive functions. The ability of lower animals to see and move, but not to speak or reason, is misleading if taken as evidence that vision and motor control are not intelligent processes on par with higher cognition. Evolution has had millions of years to compile vision and motor control into hardware and it is easy to underestimate their complexity. After all, a number of the supposed highest examples of intelligent behavior have been easiest to duplicate, such as chess playing, symbolic mathematics and logic, whereas vision and motor control have proven stubbornly difficult. Precisely the most common abilities of humans and animals seem hardest to understand, and it has been suggested that we will replace mathematicians before we replace gardeners.

Intelligent behavior requires the connection of perception to action, and it can be argued that vision and robotics will eventually assist in understanding cognition. The task of obtaining information about the environment by interpreting sensory data under noisy and uncertain conditions, and knowledge about the manipulation of objects, must have strong implications for central representations. Research in vision and robotics will also need to address higher brain functions, as we begin to ask deeper questions about problems such as the recognition and manipulation of objects, navigation through complex environments, learning of visual and motor tasks, and the control of visual attention.

2 The Study of Vision

This section describes some of the ways in which computational methods strengthen the study of biological vision. The most important contribution of the computational approach thus far has been to demonstrate just how difficult it is to solve problems in vision. Seeing is a deceptively simple task to perform. We open our eyes and suddenly capture many important aspects of the world — its structure, movement, color, texture, and so on. But hidden beneath this simple act are complex processes that transform the visual image into this rich internal description of the world.

A second contribution of the computational approach has been to show how properties of the physical world constrain the methods required to solve problems in vision. For example, the general strategies that any visual system uses to extract depth information from the two viewpoints given by the left and right eyes depends on the physics of the projection of surfaces onto the eyes and the structure of physical surfaces. The strategy used to distinguish whether a change in light intensity is due to a change in surface reflectance, surface structure or surface illumination, depends in part on the physics of light.

A third contribution of the computational approach has been to design specific *algorithms* to solve problems in vision, and to implement and test these algorithms with a computer. Such analysis forces a detailed specification of proposed methods for solution and tests the adequacy of the methods for solving visual problems. The computer implementation of vision algorithms often uncovers new aspects of a problem that were not realized in the theoretical analysis, or reveals aspects that were thought to be easy to solve but in fact turn out to be difficult. The importance of algorithms is illustrated in this section through examples of specific problems in vision.

At this stage in the study of vision, few compelling examples exist of the potentially fruitful interaction between computational studies and the experimental neurosciences. The bridge is only now being formed. Fortunately, there are some problems for which this interaction has begun to show promise. Two examples discussed in this chapter are the analysis of visual motion and detection of changes of intensity in the retinal image.[2] Section 2.2.2 describes how a computational analysis of motion measurement guided a physiological study of neurons in the middle temporal area of the extrastriate cortex. Section 2.3 shows how computational, physiological and psychophysical studies of the detection of intensity changes are together uncovering the role of some striate cortical

[2] A third problem for which there has been substantial interaction between computational and experimental studies is binocular stereopsis. Computational, perceptual and physiological studies of stereopsis are summarized in a review by Poggio and Poggio (1984).

neurons in early visual processing.

2.1 The Representational Structure of Vision

The goal of vision is to determine what is in the world and where. Biological vision must begin, however, with measurements of the amount of light reflected from surfaces in the environment onto the eye. The retinal image provided by the photoreceptors can be thought of as a large array of continuously changing numbers that represent light intensities, as shown in Figure 1. From this array of light measurements, the visual system does not achieve an understanding of what is seen in a single step. Vision proceeds in stages, with each stage producing increasingly more useful descriptions of the world. The process of vision can be viewed as the construction of a series of representations of visual information, with explicit computation that transforms one representation into the next.

It is not yet known how biological systems represent visual information, but computational studies have suggested several intermediate representations that are useful in visual processing (for example, Marr, 1982; Barrow and Tenenbaum, 1978; Horn, 1985). Representations proposed for the early stages of vision capture information that can be extracted simply and directly from the initial image. Later representions capture information that is necessary to solve complex tasks such as navigation through the environment, manipulation of objects, and recognition. Marr (1982) distinguished three representations called the Primal Sketch, the $2\frac{1}{2}$–D Sketch and the 3–D Model. The Primal Sketch is a rich description of the changes of intensity in the image, which correspond to the locations of important physical changes in the scene such as object boundaries and surface markings. The $2\frac{1}{2}$–D Sketch captures the local geometry or shape of visible surfaces in the scene, represented as the orientation or depth of surfaces at each location in the image. The 3–D Model captures the full three–dimensional structure of objects in the world, sometimes filling in hidden structure that cannot be seen. Many familiar visual processes, such as the analysis of movement, binocular stereopsis, surface shading, texture and color, can contribute to the computation of these intermediate visual representations.

Representations such as the Primal Sketch, $2\frac{1}{2}$–D Sketch and 3–D Model are tools for focusing the goals of visual computations. They make explicit what information must be computed in order to solve problems in vision. The choice of which representation to use is critical in a computational study, as some representations facilitate the solution to visual problems more than others. As an analogy, arithmetic operations such as multiplication can be carried out more easily with a representation of numbers as Arabic

a.

218	213	215	221	220	217	222	219	218	211	213	220
220	219	217	212	215	214	215	217	211	203	209	219
217	211	214	202	191	185	169	161	149	132	147	221
214	209	180	169	155	141	137	132	127	129	141	218
182	162	156	149	143	139	133	127	123	171	188	217
154	149	141	139	137	141	134	122	142	158	184	219
144	142	137	131	129	127	129	141	161	177	201	222
136	140	145	149	146	137	139	152	160	181	209	216
142	152	153	157	156	149	142	158	163	180	211	214
111	113	151	158	157	155	172	175	179	177	210	216
101	107	158	161	162	168	160	167	170	171	213	219
104	111	152	155	157	172	161	169	180	186	209	220
100	109	157	174	179	189	203	215	217	216	218	219
105	121	187	194	202	209	220	218	216	219	221	223
103	189	199	200	214	217	219	220	218	217	219	220
172	201	202	207	211	212	218	217	221	216	218	222

b.

Figure 1: The light intensities measured by a digital camera, for the rectangular area outlined in the image of (a) are shown in (b).

numerals versus Roman numerals. It is sometimes assumed in computational studies that vision proceeds sequentially from the image through the Primal Sketch, $2\frac{1}{2}$–D Sketch and 3–D Model. As vision research progresses, the relationship between these representations may become more complex.

Studies of biological vision systems have begun to examine what information is extracted from the changing retinal images. Perceptual studies address how the human system represents visual information. Examples of studies that address how the human visual system represents changes of intensity in the image are mentioned briefly in Section 2.3. At a neural level one should not expect to insert electrodes at some stage of the visual pathway and find an explicit representation such as the Primal Sketch or $2\frac{1}{2}$–D Sketch. Neurons exist that select for movement and depth, but an accurate and detailed representation of these properties may not exist explicitly in the outputs of a population of neurons.

The computational study of vision has also addressed several higher level processes, such as the control of selective visual attention, analysis of spatial relations, recognition of objects, and the organization of visual memory. While many interesting theoretical and experimental developments have emerged, the closest interaction between computational studies and the neurosciences has been in the early stages of vision. This chapter focuses on these early stages, which contribute to representations such as the Primal Sketch and $2\frac{1}{2}$–D Sketch.

2.2 Natural Constraints in Vision

An important aspect of the computational study of a visual task is to elucidate the physical assumptions necessary to solve the problem. From the changing image that reaches the eye, the human visual system derives a single, stable interpretation of what is in the scene, where it is located, and how it changes with time. For most problems that are solved in the early stages of vision, however, there is an infinity of possible solutions. To obtain a single interpretation of the image, it is necessary to make assumptions about the physical world that allow most interpretations to be ruled out, leaving one that is most plausible from a physical standpoint. The analysis of which assumptions are most appropriate for a given problem includes insights from physics, mathematics and perceptual psychology. Although less directly accessible through physiological experiments, the choice of physical assumptions constrains the type of algorithm used to solve a problem, which in turn constrains the neural mechanisms used to carry out a computation.

For the early stages of vision that precede recognition, the physical assumptions can

be general. For example, physical surfaces tend to be solid and locally rigid; points on a surface occupy a single location in space at each moment; the structure of a surface usually varies smoothly across the visual field and transforms slowly over time. Such assumptions are essential and often sufficient to solve problems such as the measurement of visual motion and the recovery of three-dimensional structure from binocular stereopsis and relative movement. This section examines the way in which physical assumptions can be used to formulate some of these problems, in order to obtain a single interpretation of the visual image.

2.2.1 The recovery of three-dimensional structure from motion

To illustrate the ambiguity that arises in the interpretation of visual information, consider the problem of deriving three-dimensional (3-D) structure from relative movement. When an object moves in space, the motions of individual points on the object differ in a way that conveys information about its 3-D structure. Suppose, for example, that the wireframe object of Figure 2a is rotated about its central vertical axis. Figure 2b shows the result of projecting this object and its movement onto the two-dimensional (2-D) image.[3] The arrows represent the projected direction and speed of movement of individual points on the object. The directions are all horizontal, but the speed of movement varies in a way that depends on the structure of the object. Wallach and O'Connell (1953) showed that the human visual system can derive the correct 3-D structure of moving objects from their changing 2-D projection alone. Other perceptual studies also demonstrated this remarkable ability (for example, Green, 1961; Braunstein, 1962, 1976; Johansson, 1973, 1975; Regan, Beverly and Cynader, 1979; Ullman, 1979).

The recovery of 3-D structure from the changing 2-D image is difficult because in theory, there are infinitely many combinations of 3-D structure and motion that could give rise to a given 2-D image. This ambiguity is illustrated with a pattern of unconnected dots in motion in Figure 3. A set of dots on the surface of a rotating transparent cylinder are projected onto a 2-D display screen, using an orthographic projection (Figure 3a). A birds' eye view of this projection is shown in Figure 3b. When the dots are projected onto the image, information about their location and movement in depth is lost. Yet when human subjects view only the 2-D pattern of moving dots, they derive a vivid impression of the dots lying on a transparent cylinder in rotation. Clearly, many interpretations are possible. The dots actually lie on the flat plane of the display screen, but in principle could lie anywhere in depth and undergo any movement

[3]For simplicity, an orthographic projection is used, in which points in space are projected in parallel and in the direction perpendicular to the image plane.

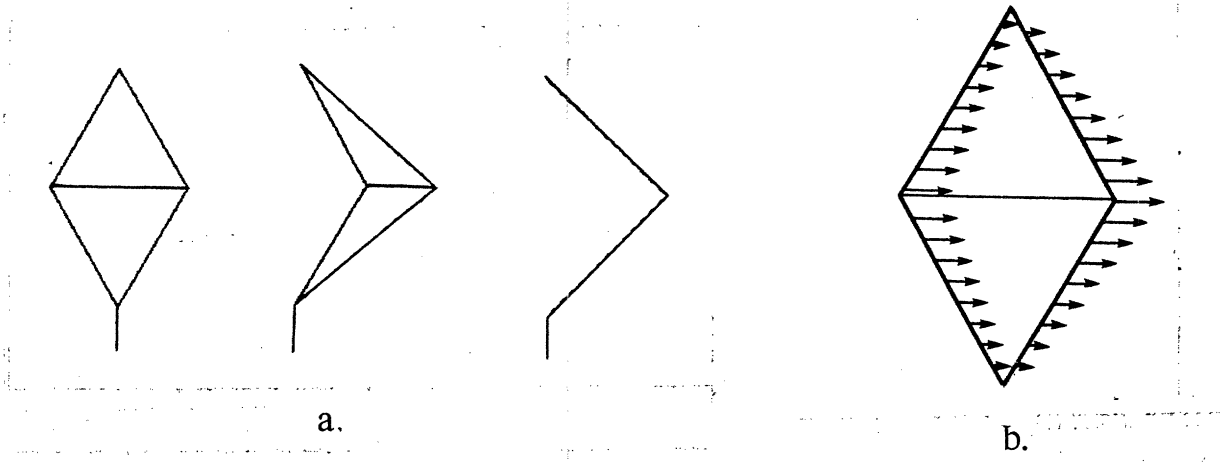


Figure 2: Deriving three-dimensional structure from two-dimensional motion. (a) Three views of a 3-D wireframe object that is rotating about a central vertical axis. (b) The projected 2-D image and motion of the object. The arrows represent the projected 2-D velocity of individual points on the object.

in depth. The random field of moving dots shown from a birds' eye view in Figure 3c also gives rise to the same projected 2-D image. How does the human visual system conclude that the moving dots lie on the surface of a rotating cylinder?

Computational studies have used the assumption of *rigidity* to derive structure from motion. These studies assume that if it is possible to interpret a changing 2-D image as the projection of a rigid 3-D object in motion, then such an interpretation should be chosen (Ullman, 1979, 1983; Clocksin, 1980; Prazdny, 1980, 1983; Longuet-Higgins, 1981; Longuet-Higgins and Prazdny, 1981; Tsai and Huang, 1981; Hoffman and Flinchbaugh, 1982; Bobick, 1983). When the rigidity assumption is used in this way, the recovery of structure from motion requires the computation of the rigid 3-D objects that would project onto a given 2-D image. The rigidity assumption was suggested by perceptual studies that described a tendency for the human visual system to choose a rigid interpretation of moving elements (Wallach and O'Connell, 1953; Gibson and Gibson, 1957; Green, 1961; Johansson, 1975).

Computational studies have shown that it is possible to use the rigidity assumption to derive a unique 3-D structure from a changing 2-D image. Furthermore, it is possible to derive this unique 3-D interpretation by integrating image information only over a limited extent in space and in time. For example, suppose that a rigid object in motion is projected onto the image using the orthographic projection illustrated in Figure 3. Three distinct views of four points on the moving object are sufficient to compute a unique rigid 3-D structure for the points (Ullman, 1979). In general, if only two views

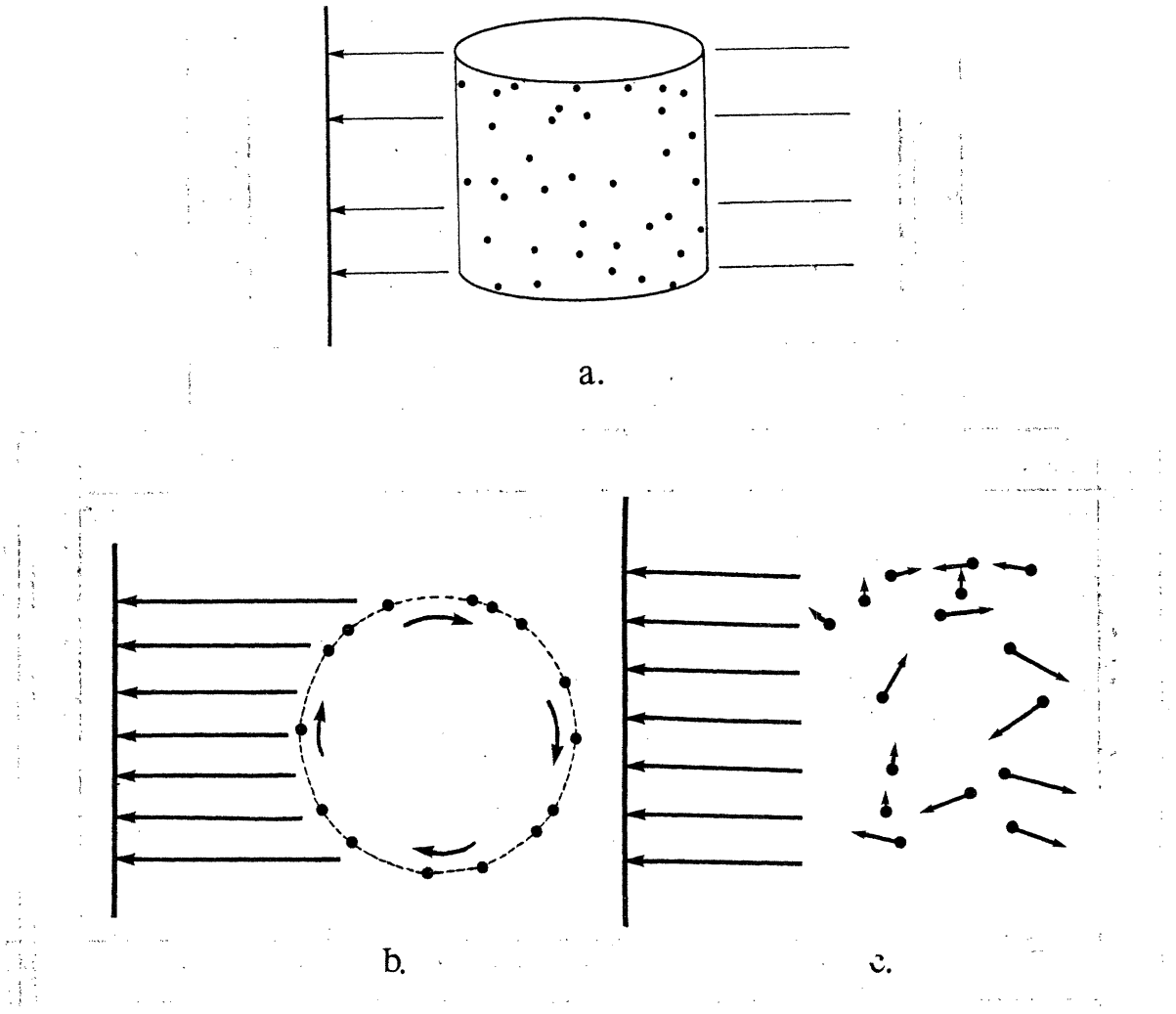


Figure 3: The ambiguity of interpreting structure from motion. (a) A set of dots on the surface of a rotating transparent cylinder are projected onto a 2-D display screen. (b) Birds' eye view of the projection of the dots in (a). (c) A field of randomly moving dots that project to the same 2-D image as the dots shown in (b).

of the moving points are considered or fewer points are observed, there are multiple rigid 3–D structures consistent with the changing 2–D projection. Suppose that a perspective projection of objects onto the image is used instead. In this case, two distinct views of 7 points in motion are usually sufficient to compute a unique 3–D structure for the points (Tsai and Huang, 1981). Other theoretical results regarding the recovery of structure from motion are summarized in Ullman (1983). These theoretical results are important for two reasons. First, they show that by using the rigidity assumption, it is possible to recover a unique structure from motion information alone. Second, they show that it is possible to recover this structure by integrating image information over a small extent in space and in time. The second observation could bear on the neural mechanisms that compute structure from motion — in principle, they need only integrate motion information over a limited area of the visual field and a limited extent in time.

Computational studies of the recovery of structure from motion also provide algorithms for deriving the structure of moving objects (for example, Ullman, 1979; Longuet–Higgins, 1981; Tsai and Huang, 1981). Typically, measurements of the positions or velocities of features in the image give rise to a set of mathematical equations whose solution represents the desired 3–D structure. The algorithms generally derive this structure by using motion information that is extracted over a limited area of the image and a limited extent in time. Testing of these algorithms reveals that although this strategy is possible in theory, it is not reliable in practice. A small amount of error in the image measurements can lead to very different (and often incorrect) 3–D structures (Ullman, 1983, 1984). This suggests that an algorithm for deriving structure should use image information that is more extended in space or time or both. Perceptual studies have indicated that the human visual system also requires an extended time period to reach an accurate perception of 3–D structure (Wallach and O'Connell, 1953; White and Mueser, 1960; Green, 1961).

Most methods for recovering structure from motion compute a 3–D structure only when it is possible to interpret the changing image as the projection of a rigid object in motion. They otherwise either yield no interpretation of structure or yield a solution that is incorrect or unstable. Yet the human visual system can derive some sense of structure for nonrigid objects in motion (Johansson, 1964, 1978; Jansson and Johansson, 1973). Furthermore, displays of rigid objects in motion sometimes give rise to the perception of a somewhat distorting object (Wallach, Weisz and Adams, 1956; White and Mueser, 1960; Green, 1961; Braunstein, 1962; Hildreth, 1984). These observations suggest that while the human visual system tends to choose rigid interpretations of a changing image, it probably does not use the rigidity assumption in the strict way that

previous computational studies have suggested.

Ullman (1984) proposed a more flexible method for deriving structure from motion that allows both rigid and nonrigid motion to take place. It makes use of the rigidity assumption, but in a different way from previous studies. The algorithm maintains an internal model of the structure of a moving object, which consists of the estimated 3–D coordinates of points on the object. The model is continually updated as new positions of image features are considered. Initially, it is assumed that the object is flat, if no other cues to 3–D structure are present. Otherwise, its initial structure may be determined by other cues available, from binocular stereopsis, shading, texture or perspective. As each new view of the moving object appears, the algorithm computes a new set of 3–D coordinates for points on the object. In particular, the algorithm chooses a new set of coordinates that maximize the rigidity in the transformation from the current model to the new positions. This is achieved by minimizimg the change in the 3–D distances between points in the model. Thus the algorithm interprets the changing 2–D image as the projection of a moving 3–D object that changes as little as possible from one moment to the next. Through a process of repeatedly considering new views of objects in motion and updating the current model of their structure, the algorithm builds up and maintains a 3–D model of the objects. If objects deform over time, the 3–D model computed by the algorithm also changes over time.

The method proposed by Ullman (1984) for recovering structure from motion was motivated in part by the limitations of previous computer algorithms and in part by knowledge of the human visual system. The method has overcome the limitations of previous computational studies in two ways. First, it provides a reliable recovery of structure in the presence of error in the image measurements, by integrating image information over an extended time period. Second, it allows the interpretation of non-rigid motions. These are essential qualities for any method that is proposed as a viable model for the recovery of structure from motion by the human visual system. This method also has other attributes that are consistent with human perceptual behavior: (1) it sometimes yields a nonrigid interpretation of rigid structures in motion, (2) a brief viewing time results in a structure that is "flatter" than the true structure of the object, (3) it allows a 3–D interpretation of scenes containing as few as two points in motion (Borjesson and von Hofsten, 1973; Johansson, 1975), and (4) it provides a natural means for integrating multiple sources of 3–D information. The existence of a detailed model for recovering structure also allows predictions that could form the basis for further psychophysical experiments. For example, computer experimentation with this method shows that the recovery of the structure of rotating objects degrades as their axis of

rotation is tilted away from the plane of the image (Grzywacz and Hildreth, 1985). This raises the question of whether the ability of the human visual system to recover the structure of rotating objects varies with the orientation of the axis of rotation in space.

This discussion of the structure–from–motion problem illustrates a number of important points that also arise in the computational study of other problems in the early stages of vision. First, a single solution to the problem cannot be obtained from information in the image alone; additional constraint is required. This is a general aspect of vision problems that makes them especially difficult to solve. Second, physics and mathematics can be used to show that a general physical assumption such as rigidity is sufficient to solve the structure–from–motion problem uniquely. Third, an assumption such as rigidity can be incorporated in many ways into an algorithm to recover structure. The development of a reliable algorithm requires a cycling between computer implementation, testing and refinement. Finally, perceptual studies can suggest and test particular assumptions and reveal aspects of the algorithm used by the human visual system for solving a given problem. It is typical of computational studies that the initial methods proposed for solving a problem only loosely consider the detailed observations of biological systems. These first studies uncover useful aspects of the problems. Later studies then combine this knowledge of the problem with observations of biological systems to derive models that more closely mimic the computations carried out in biological systems.

To study the neural mechanisms that underly the recovery of structure from motion, it would be useful to explore the properties of neurons that respond selectively to the interpreted position or movement in depth of features in monocularly viewed changing patterns such as those illustrated in Figures 2 and 3. There exist neurons in area 18 of the cat visual cortex (Cynader and Regan, 1978, 1982) and area V1[4] of the primate visual cortex (Poggio and Talbot, 1981) that appear to be selective for direction of movement in depth. These studies used binocularly viewed moving bars, however, so they address the interaction between binocular stereopsis and motion measurement for the recovery of movement in three dimensions, rather than the recovery of structure from motion alone.

2.2.2 The measurement of visual motion

As a second example of the ambiguity that arises in the interpretation of visual information, we examine the problem of measuring movement in the changing 2–D image.

[4]Area V1 is also referred to as area 17, striate cortex, or primary visual cortex.

Consider the computation of the projected 2-D velocity field illustrated in Figure 2b. Suppose that the movement of features in the image was first detected using operations that examine only a limited area of the image. For example, movement might be detected by neural mechanisms with spatially limited receptive fields. Such mechanisms can provide only partial information about the true motion of features in the image. This is a consequence of the *aperture problem* illustrated in Figure 4a (Wallach, 1976; Fennema and Thompson, 1979; Burt and Sperling, 1981; Horn and Schunck, 1981; Marr and Ullman, 1981; Adelson and Movshon, 1982). Suppose that an extended feature such as the edge **E** moves across the image, and that its movement is observed through a window defined by the circular aperture **A**. Through this window, it is only possible to observe the movement of the edge in the direction perpendicular to its orientation. The component of motion along the orientation of the edge is invisible through this limited aperture. Thus it is not possible to distinguish between motions in the directions **b**, **c** and **d**. This property is true of any motion detection operation that examines only a limited area of the image. Neural movement detectors with spatially limited receptive fields, for example, can directly measure only the component of motion in the direction perpendicular to the orientation of moving image features.

As a consequence of the aperture problem, the measurement of motion in the changing image requires two stages of analysis: the first stage measures components of motion in the direction perpendicular to image features; the second combines these components of motion to compute the full 2–D pattern of movement in the image. In Figure 4b, a circle undergoes pure translation to the right. The arrows along the contour represent the perpendicular components of velocity that can be measured directly from the changing image. These component measurements each provide some constraint on the possible motion of the circle. Its true motion, however, can be determined only by combining the constraints imposed by these component measurements. The movement of some features such as corners or small specks can be measured directly. In general, however, the first measurements of movement provide only partial information about the true movement of features in the image.

The measurement of movement is difficult because in theory, there are infinitely many patterns of motion that are consistent with a given changing image. For example, in Figure 4c, the contour **C** rotates, translates and deforms to yield the contour **C′** at some later time. The true motion of the point **p** is ambiguous. Additional constraint is required to identify a single pattern of motion. Many physical assumptions could provide this additional constraint. One possibility is the assumption of pure translation. That is, it is assumed that velocity is constant over small areas of the image. This

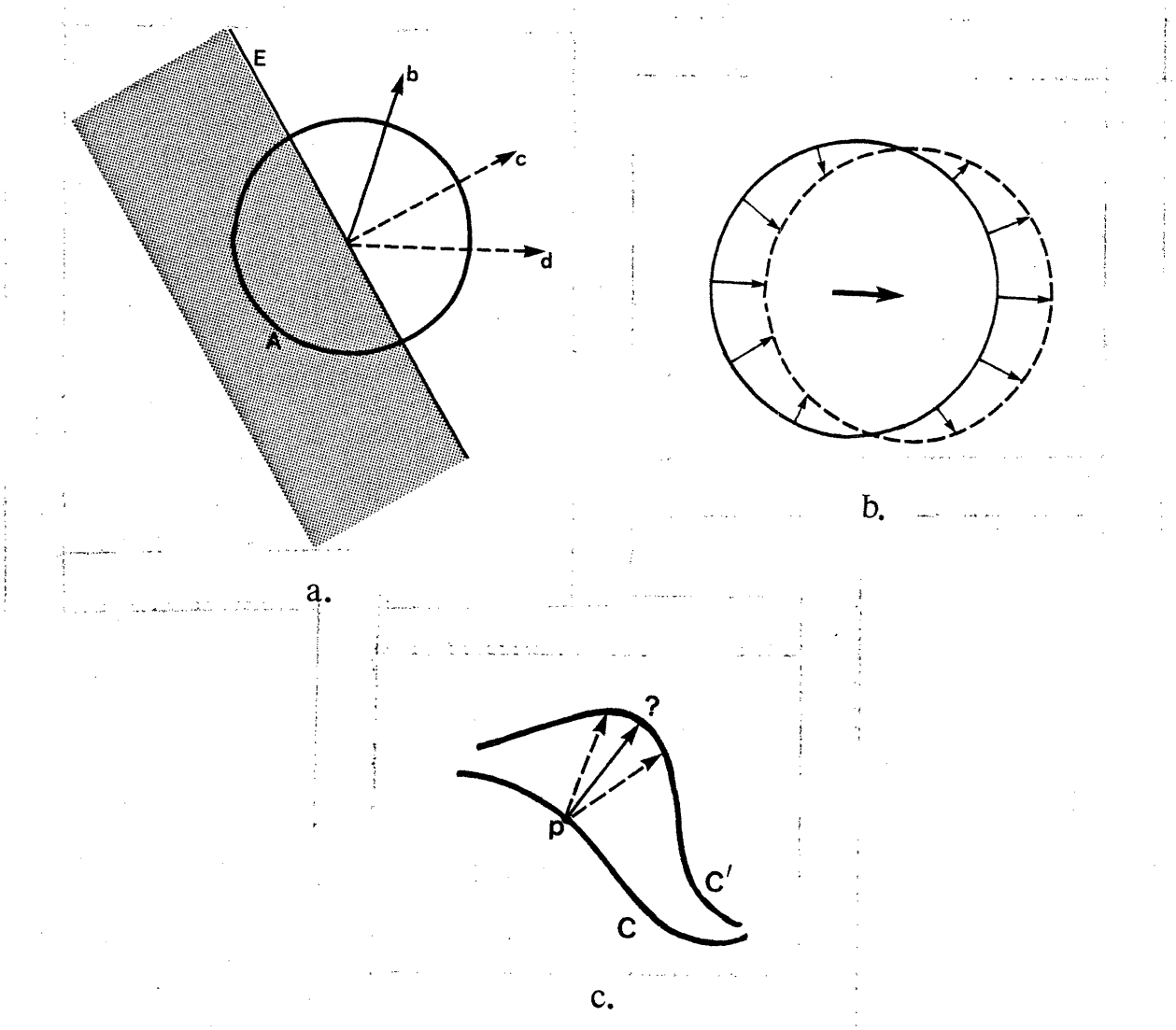


Figure 4: The aperture problem. (a) A motion detector that views the moving edge **E** through a limited aperture **A** detects only the component of motion **c** in the direction perpendicular to the edge. (b) A circle undergoing pure translation to the right. The arrows along the contour represent the perpendicular components of velocity obtained from the changing image. (c) The contour **C** undergoes translation, rotation and deformation to yield the contour **C'** at some time later. The true motion of the point **p** is ambiguous.

assumption has been used both in computer vision studies and in biological models of motion measurement (for example, Lappin and Bell, 1976; Pantle and Picciano, 1976; Fennema and Thompson, 1979; Anstis, 1980; Marr and Ullman, 1981; Thompson and Barnard, 1981; Adelson and Movshon, 1982; Lawton, 1983). Methods that assume pure translation may be used to detect sudden movements or to track objects across the visual field. These tasks may require only a rough estimate of the overall translation of objects across the image. Tasks such as the recovery of 3-D structure from motion require a more detailed measurement of relative motion in the image. The analysis of variations in motion such as those illustrated in Figure 2b requires the use of a more general physical assumption.

Recent computational studies have assumed that velocity varies smoothly across the image (Horn and Schunck, 1981; Hildreth, 1984; Nagel, 1984). The assumption rests on the principle that physical surfaces are generally smooth. Variations in the structure of a surface are usually small, compared with the distance of the surface from the viewer. When surfaces move, nearby points tend to move with similar velocities. There exist discontinuities in movement at object boundaries, but most of the image is the projection of relatively smooth surfaces. Thus, it is natural to assume that image velocities vary smoothly over most of the visual field. A unique pattern of movement can be obtained by computing a velocity field that is consistent with the changing image and has the least amount of variation possible. In other words, a pattern of movement is derived, for which nearby points in the image move with velocities that are as similar as possible.

The use of the smoothness assumption for motion measurement has several important attributes from a computational perspective. First, it allows general motion to be analyzed. Surfaces can be rigid or nonrigid, undergoing any movement in space. It is always possible to compute a projected velocity field that preserves the real variation in the local pattern of movement. Second, the smoothness assumption can be embodied into the motion measurement computation in a way that guarantees a unique solution (Hildreth, 1984). Third, the velocity field of least variation can be computed straightforwardly, using standard computer algorithms (Horn and Schunck, 1981; Hildreth, 1984).

From the perspective of perceptual psychology, one can ask whether the human visual system derives patterns of movement that are consistent with those predicted by a computation that uses the smoothness assumption. In particular, one can ask whether an incorrect pattern of motion is perceived in situations where a computer algorithm also fails. The method for computing the velocity field suggested by Hildreth

(1984) is guaranteed to yield the correct solution for at least two classes of motion: (1) pure translation, and (2) general motion (translation and rotation) of rigid 3–D objects whose edges are essentially straight. For example, the computation yields the correct velocity field for the moving objects of Figures 2a and 4b. For the case of smooth curves undergoing rotation, this computation sometimes yields a solution that differs from the correct projected velocity field. The human visual system also appears to derive an incorrect perception of motion in these situations. Three examples are shown in Figure 5. The true velocity fields for these moving figures are shown in Figures 5a, 5c, and 5e. The short line segments along the smooth contours represent true directions and speeds of movement of individual points on the contours. The velocity fields of least variation that are consistent with the changing images are shown in Figures 5b, 5d and 5f. The first example is a logarithmic spiral whose image rotates about its center. Human observers perceive an expansion or contraction of a rotating spiral, depending on its direction of motion (Holland, 1965). Thus, the true motion is pure rotation, but the perceived motion contains a large radial component. Consistent with this perception, there is a large radial component in the smoothest velocity field shown in Figure 5b, particularly toward the center of the spiral. The second figure is an ellipse that is almost circular and rotating about its center. Wallach, Weisz and Adams (1956) showed that human observers do not perceive the rotation of the ellipse; rather, they perceive the major and minor axes of the ellipse as pulsating inward and outward. This perception is also consistent with the smoothest velocity field shown in Figure 5d. Finally, if a deformed circle such as that shown in Figure 5e is rotated about its center, the circular part of the figure appears to stand still, while the bump travels around the perimeter (Wallach, Weisz and Adams, 1956), consistent with the smoothest velocity field shown in Figure 5f. Many other examples exist of the consistency of human motion perception with a computation that embodies the smoothness assumption (Hildreth, 1984).

The motion measurement problem can also be examined from a physiological perspective. Early movement detectors in biological systems have spatially limited receptive fields and therefore face the aperture problem. Stimulated by a theoretical analysis of the aperture problem, Movshon *et al.* (1985) sought and found direct physiological evidence for a two–stage motion measurement computation in the primate visual system. Two visual areas that include an abundance of motion–sensitive neurons are cortical areas V1 and MT.[5] The experiments of Movshon *et al.* (1985) indicated that the selectivity of neurons in area V1 for direction of movement is such that they could

[5]MT is the middle temporal area of the extrastriate cortex, located in the posterior bank of the superior temporal sulcus (STS).

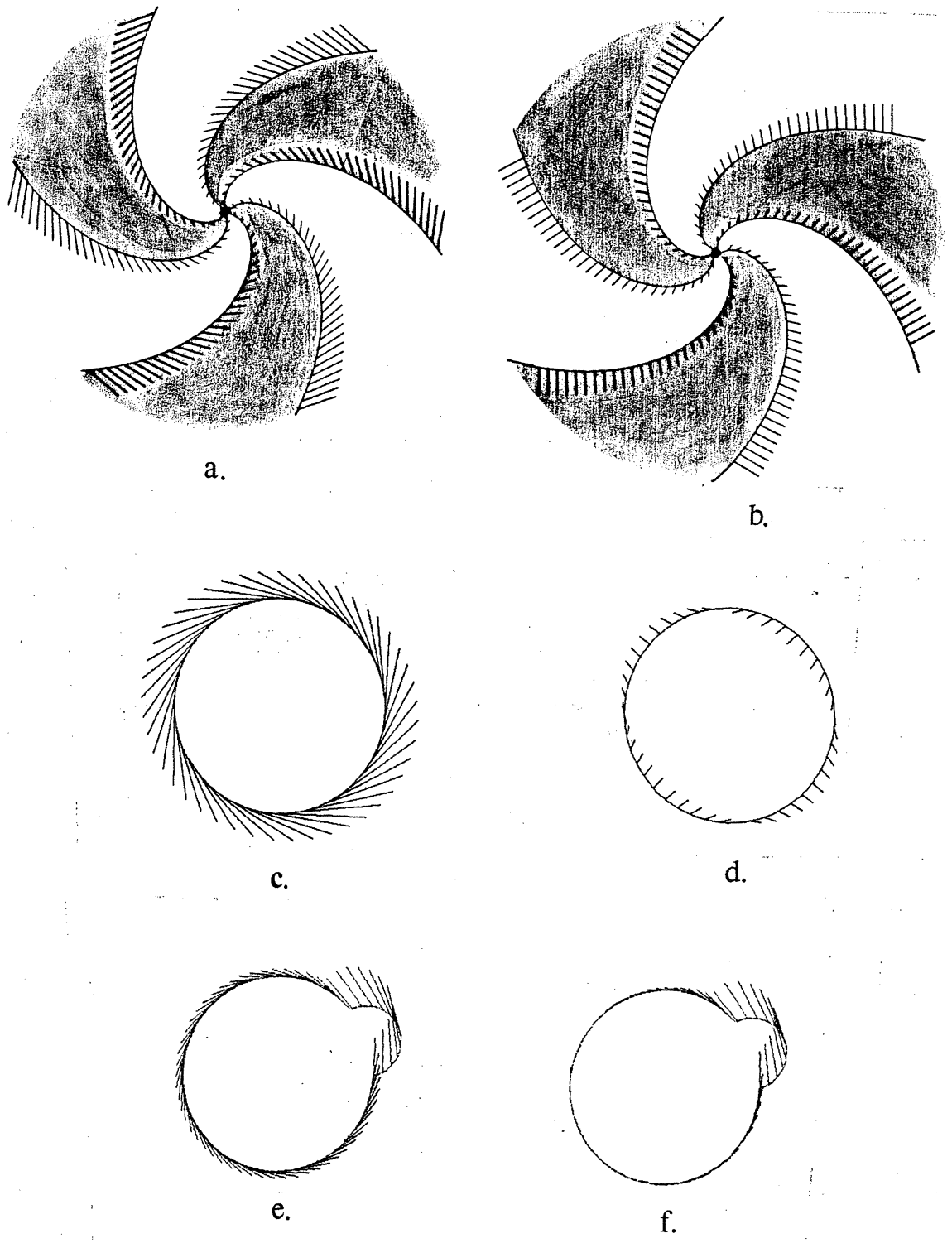


Figure 5: Motion illusions. (a), (c) and (e) The true velocity fields for a logarithmic spiral, ellipse and deformed circle, respectively, rotating about their centers. The short line segments along the smooth contours represent the direction and speed of movement of individual points on the contours. (b), (d) and (f) The smoothest velocity fields that are consistent with the rotating patterns shown in (a), (c) and (e), respectively.

only provide the component of motion in the direction perpendicular to the orientation of image features. Area MT, however, contains a subpopulation of cells, referred to as *pattern* cells, that appear to combine motion components to compute the real 2–D direction of velocity of a moving pattern. This study used visual stimuli that consist of superimposed sinewave gratings of different orientations, each moving in the direction perpendicular to its orientation. These experiments do not yet distinguish between the use of the simple assumption of pure translation, as suggested in the study (Movshon *et al.*, 1985), versus the more general smoothness assumption. Stimulus patterns undergoing more complicated motions are required to make such a distinction. If the pattern cells in area MT embody the assumption of smoothness in their computation of motion, one would expect to find direct interaction between pattern cells that analyze nearby areas of the visual field.

The study of Movshon *et al.* (1985) illustrates the importance of integrating theoretical and experimental studies. Theoretical studies of motion measurement showed that a particular type of computation should take place in order to solve this problem, namely, the combination of perpendicular components of motion to determine the real direction of motion of a pattern in two dimensions. This observation then led to a specific physiological study aimed at determining where in the visual pathway this computation takes place.

Poggio and Koch (1984) presented a hypothetical neural implementation of the computation of the smoothest velocity field that uses known properties of neural hardware. Poggio and Koch first designed electrical and chemical networks to perform this computation in an analog manner. From these networks, a neural circuit was then designed that behaves in a similar way. Examples of the electrical and neural networks are shown in Figure 6. In the network of Figure 6a, the currents $\mathbf{I}_i$ and conductances $\mathbf{g}$ and $\mathbf{g}_i$ represent measurements of the perpendicular components of velocity and other properties of a moving contour obtained directly from the image. The voltages $\mathbf{V}_i$ represent the tangential component of velocity[6] that is recovered by the computation of the full 2–D velocity field. These analog networks allow a fast computation of the smoothest velocity field. In the corresponding neural implementation of Figure 6b, the tangential component of the velocity field is represented by the voltages $\mathbf{V}_i$ along a dendrite, which are sampled by dendro–dendritic synapses. Measurements from the image are represented by synaptically mediated current injections $\mathbf{I}_i$ and other synaptic inputs $\mathbf{R}_i$ that control the membrane resistance. The full 2–D velocity field is represented implicitly

[6]The tangential component is the component of velocity in the direction parallel to the orientation of features in the image.

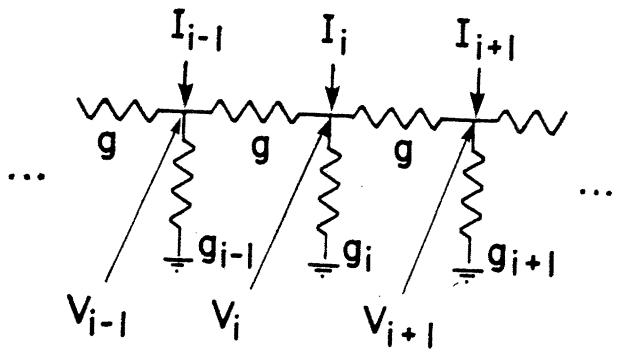
by the combination of the currents $\mathbf{I}_i$ and the voltages $\mathbf{V}_i$. This hypothetical neural implementation was not intended as a specific model for the measurement of motion in the area MT. Rather, its intent was to show that it is possible for neural hardware to exploit a model of this computation that incorporates a general assumption such as smoothness of the velocity field. Models such as this can help to focus experimental questions regarding the actual neural circuitry in areas such as MT.

The assumption of smoothness of physical surfaces and their motion is not always appropriate. Although much of an image can represent the projection of relatively smooth surfaces, sudden changes or discontinuities may exist in surface structure and motion, both within objects and at object boundaries. The detection of discontinuities in motion is an important problem that must be considered together with computations of motion that rely on the smoothness assumption.

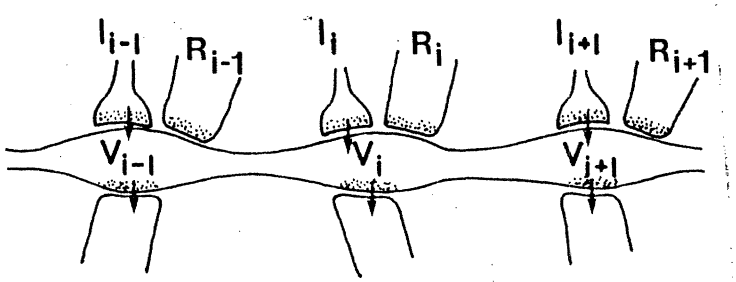
This discussion of the measurement of motion again illustrates a number of important aspects of the computational study of vision. Similar to the recovery of structure from motion, a unique pattern of movement cannot be obtained from information in the changing image alone. This problem requires additional constraint that is imposed by properties of the physical world. The need to relate vision to properties of the external world is not a new idea. Gibson (1950, 1966, 1979) argued this point forcibly in his theories of vision. Computational studies have taken this observation further. A full understanding of how the human visual system solves a problem in vision must make explicit these additional assumptions, their physical justification, and how they can be incorporated into a specific computation in a way that yields a unique and stable solution to the problem. The design and implementation of algorithms that embody a particular assumption provides a useful tool for making predictions from the computational model that can be tested directly through perceptual experiments. For many visual processes it is difficult to predict the outcome of a computational model without an algorithm that implements the model. Finally, theoretical studies reveal the computations that must take place in order to solve problems such as the measurement of motion, which can guide physiological studies that explore where these computations take place in the visual pathway. The study of Movshon *et al.* (1985) is one example of this interaction.

2.3 From Theory to Implementation: the Detection of Intensity Changes

This section examines the role of computational, physiological and psychophysical approaches in the study of vision, through the problem of detecting changes of intensity



a.



b.

Figure 6: Analog models of the velocity field computation. (a) A simple resistive network that computes the smoothest velocity field. The conductances g and g_i, and the currents I_i represent properties of a moving contour that are measured directly from the image. The 2-D velocity field along the contour is represented implicitly by the combination of these inputs and the resulting voltages V_i. (b) A hypothetical neural implementation of the circuit shown in (a). Synaptic mediated currents I_i, and additional inputs R_i represent properties of a moving contour. The resulting voltages V_i, sampled by dendro–dendritic synapses, together with the input currents, represent local velocities along the contour.

in the visual image. Important physical features such as object boundaries, surface markings, shadows and surface textures, give rise to changes in the light intensity that is reflected onto the eye. The detection of these intensity changes in the image provides the first clue about the structure of the scene and is considered an important aspect of early visual processing. A description of intensity changes is also useful for the subsequent analysis of motion, binocular stereopsis, texture and other visual properties. Until the late 1970's, largely independent investigations of the early stages of vision took place in physiology, psychophysics and computer vision. Recent studies have integrated the findings of these three fields in a way that both reveals the computations necessary to detect intensity changes and contributes understandings about the function of neurons in the visual pathway.

Early studies in computer vision made a number of important observations concerning the detection of intensity changes, or edges as they are often called (for reviews see Davis, 1975; Pratt, 1978; Hildreth, 1983; Horn, 1985). First, in real images, intensity typically changes from one location in the image to the next and not all of these changes are due to significant physical events. Some, for example, are due to noise in the sensors. If the intensity measurements are smoothed, however, minor fluctuations of intensity can be removed, leaving only the most important. Second, the detection and localization of intensity changes can be facilitated by performing a first or second derivative [7] operation on the smoothed intensities. These smoothing and derivative operations are illustrated in Figure 7. Figure 7a shows a one-dimensional intensity profile that represents the intensity of light measured along a horizontal line in a natural image. These intensities are then smoothed in Figure 7b. Spatial changes in the smoothed intensity profile give rise to peaks in the first derivative shown in Figure 7c, or zero-crossings (transitions between positive and negative values) in the second derivative shown in Figure 7d. This can be seen by following the dotted lines from Figure 7b through 7d. These peaks and zero-crossings are easy to detect, and properties such as the position and height of the peaks can be used to compute the location, sharpness and contrast of the intensity changes in the image. Properties of the intensity changes in turn provide useful information about the underlying physical changes in the scene, although little is known at this time about how this interpretation of intensity changes might proceed. A third observation from early computational studies is that important changes in the image occur at different spatial resolutions and can often be detected by smoothing the image by different amounts.

[7] The first derivative of a function is a measure of the rate of change of the function and the second derivative is the rate of change of the first derivative.

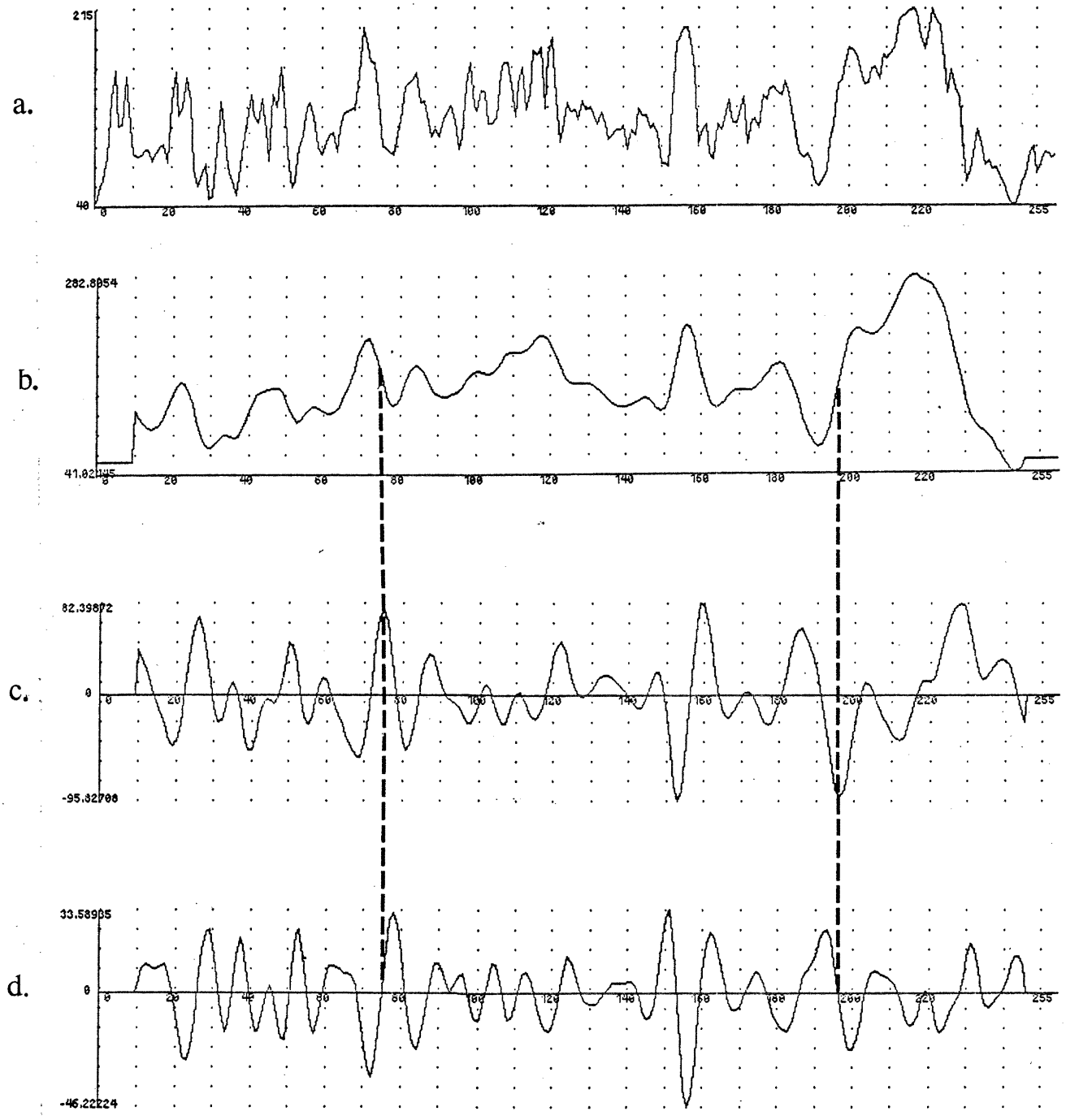


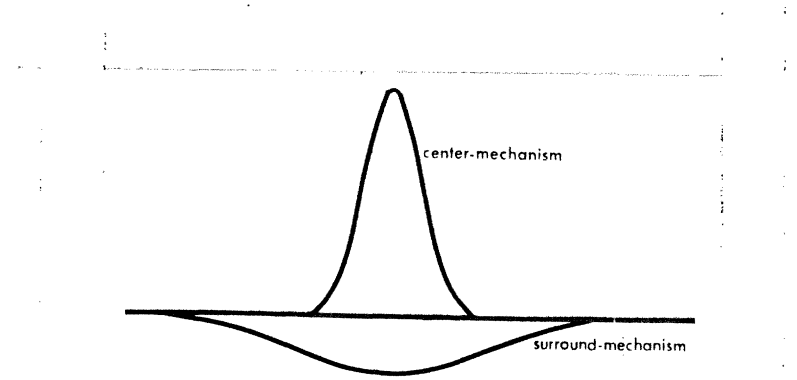
Figure 7: Detecting intensity changes. (a) One-dimensional intensity profile that represents the light intensities measured along a horizontal line of a natural image. (b) The result of smoothing the intensity profile shown in (a). (c) The first derivative of the smoothed intensity profile shown in (b). (d) The second derivative of the smoothed intensity profile shown in (b). The dotted lines show the relationship between significant changes in (b), peaks in (c) and zero-crossings in (d).

Physiological studies suggest that the analysis of intensity changes may be one of the first stages of processing in biological vision systems. Early electrophysiological recordings showed that retinal ganglion cells have a spatial receptive field with an antagonistic center–surround organization (Kuffler, 1953), whose shape can be described as the difference of two Gaussian functions, shown in Figure 8a (Rodieck and Stone, 1965; Enroth–Cugell and Robson, 1966). These early studies also distinguished ON and OFF center cells, shown in Figure 8b. In the case of ON center cells, light in the center of the receptive field increases the cell's response, while light in the surround decreases the cell's response. OFF center cells behave in the opposite manner. Rodieck (1965) described the output of the retinal ganglion cells as the convolution[8] of the changing image with the spatial difference–of–Gaussians (DOG) function, combined with a particular temporal filtering. This spatial filtering with the DOG function enhances changes in light intensity.

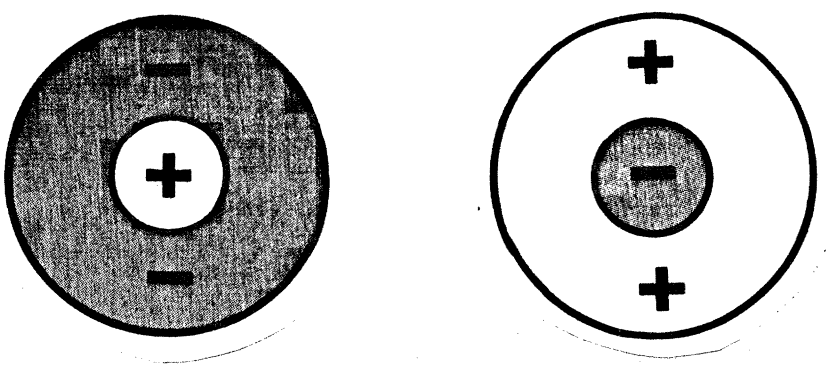
Physiological studies also have revealed the existence of different classes of retinal ganglion cells. The two main cell types have been labelled X and Y cells in the cat[9] (Enroth–Cugell and Robson, 1966; Cleland, Dubin and Levick, 1971). X cells generally have smaller receptive fields than Y cells (*cat:* Enroth–Cugell and Robson, 1966; Boycott and Wassle, 1974; Peichl and Wassle, 1979; *monkey:* deMonasterio and Gouras, 1975), X cells sum their inputs linearly, while Y cells are highly nonlinear (*cat:* Enroth–Cugell and Robson, 1966; Hochstein and Shapley, 1976; *monkey:* Schiller and Malpeli, 1977; deMonasterio 1978a), and X cells exhibit color selectivity while Y cells generally respond to a broad range of colors (*monkey:* deMonasterio and Gouras, 1975; Schiller and Malpeli, 1977; deMonasterio, 1978b). Finally, X cells respond in a sustained manner to temporal changes in light intensity, while Y cells respond in a transient manner (*cat:* Cleland, Dubin and Levick, 1971; Cleland, Levick and Sanderson, 1973; *monkey:* deMonasterio, 1978a). The optic nerve carries the output of the X and Y retinal ganglion cells to the lateral geniculate nucleus (LGN), where the main properties of these two systems of cells are largely preserved (Cleland, Dubin and Levick, 1971; Hoffman, Stone and Sherman, 1972; Dreher and Sanderson, 1973). The output of the LGN then forms one of the main sources of input to area V1 of the visual cortex. With regard to function, it has been proposed that the X system plays a greater role in the spatial analysis of the image, while the Y system serves to analyze movement or temporal change (for example,

[8] Convolution is an operation that weighs inputs within some area of the image by different amounts and sums the results.

[9] The X and Y cell distinction only strictly applies to the cat, but cell classes with similar properties exist in the monkey, so studies of retinal ganglion cells in the monkey are also listed here.



a.



b.

Figure 8: The receptive fields of retinal ganglion cells. (a) The shape of the spatial receptive fields of retinal ganglion cells, described quantitatively as a difference of two Gaussian functions, a narrow positive one and broader negative one. (b) ON and OFF center cells, which respond in an opposite manner to light stimulation in the central and surrounding areas of their receptive fields.

Tolhurst, 1973; Kulikowski and Tolhurst, 1973; Ikeda and Wright, 1972, 1975).

Early recordings in the visual cortex of cat and monkey revealed cells that respond vigorously when simple features such as edges or bars of a particular orientation and contrast move across the visual field (Hubel and Weisel, 1962, 1968; Pettigrew, Nikara and Bishop, 1968; Bishop, Coombs and Henry, 1971; Goodwin, Henry and Bishop, 1975). Cortical cells also segregate into different classes on the basis of physiological properties. Hubel and Wiesel (1968) originally distinguished four functional classes, labelled nonoriented, simple, complex and hypercomplex. The main class of interest here are the simple cells, which respond optimally to an edge or bar of a particular orientation moving across their receptive field. Some simple cells are also selective for the sign of contrast of the edge or bar and its direction of motion. In a quantitative study of cortical cells in the rhesus monkey, Schiller, Finlay and Volman (1976) further subdivided simple cells into seven distinct classes, on the basis of the spatiotemporal distribution of their response to moving edges and rectangles and stationary flashed stimuli. With regard to the function of cortical cells, it was suggested by Barlow (1972) and others that these cells may be the neural correlates of primitive feature detectors.

Perceptual studies also stressed the importance of intensity changes in early visual analysis. As early as 1865, Mach observed that our perceptual system is particularly sensitive to and actually enhances spatial changes in light intensity. Studies by Cornsweet (1970), Land (1959a,b; Land and McCann, 1971) and others also revealed that sharp changes of intensity play an important role in the perception of lightness, while gradual changes are essentially ignored.

A second important psychological discovery is that the visual system initially processes the image through a number of separate channels that differ in the way they analyze spatial and temporal variations of intensity (for example, Campbell and Robson, 1968; Blakemore and Campbell, 1969; Graham and Nachmias, 1971; Kulikowski and Tolhurst, 1973; Tolhurst, 1973, 1975; Spitzberg and Richards, 1975; Breitmeyer and Ganz, 1977; Cowan, 1977; Graham, 1977; Watson and Nachmias, 1977; Wilson and Bergen, 1979). Some channels are more sensitive to slower spatial variations of intensity in the image, while other channels are more sensitive to rapid fluctuations. The channels also differ in their sensitivity to temporal variations of intensity. Wilson and Bergen (1979) proposed a quantitative model of the operations performed by these channels, which incorporates a spatial filtering of the image with the DOG function found in physiological studies.

To illustrate how theoretical, physiological and psychophysical studies each contribute toward the computations that underly visual processing, we examine here a par-

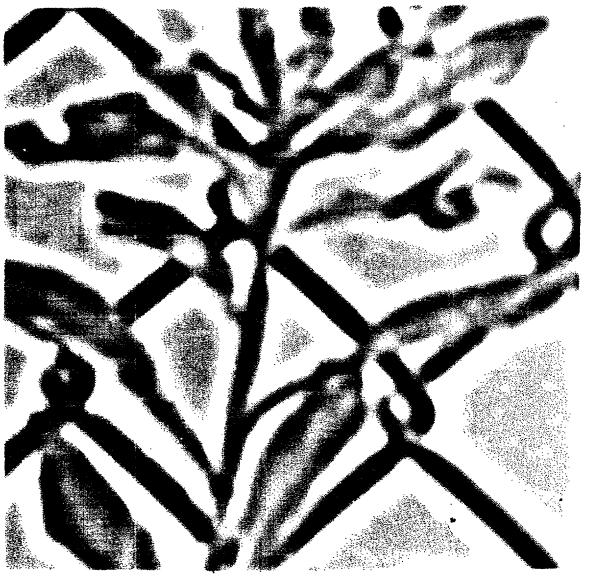
ticular method for detecting intensity changes proposed by Marr and Hildreth (1980). The example was chosen for several reasons. First, the method grew out of computational arguments and integrated a number of the important ideas that had been developed in earlier studies of edge detection. Second, it has stimulated the design of neural models to implement the computations. Third, it has motivated physiological and psychophysical experimentation aimed at testing its validity as a model of a stage of processing in biological vision systems. At this time it remains only a hypothesis for one aspect of early vision.

Marr and Hildreth (1980) first proposed on theoretical grounds that to detect intensity changes, the image should be filtered with an operator whose spatial shape is given by the Laplacian operator applied to a Gaussian distribution, which is closely approximated by the DOG function. This filtering embodies operations that were considered important in early edge detection studies. The spatial extent of the DOG function serves to smooth the image and the center–surround mechanism performs a kind of second derivative operation. Figure 9 shows an example of the result of this filtering computation. The image of Figure 9a is shown filtered through a DOG function in Figure 9b. The filtered image contains positive and negative values, with the most positive shown in white and most negative in black. The ON and OFF center retinal ganglion X cells can be thought of as carrying the positive and negative parts of this DOG–filtered image. When viewing the image of Figure 9a, the ON center cells are expected to be most active in the brighter areas of the image of Figure 9b, and the OFF center cells most active in the darker areas.

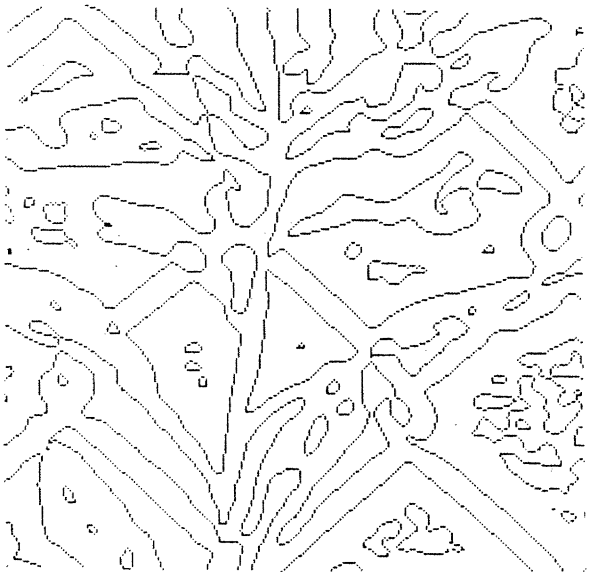
Marr and Poggio (1979) observed that the elements in the output of the filtering stage, which correspond to the locations of significant intensity changes in the image, are the zero–crossings mentioned earlier. These zero–crossings are the contours that separate the positive and negative regions of the output of the filtering stage. The zero–crossings of the filtered image of Figure 9b are shown in Figure 9c. In addition to the position of the zero–crossings, one also can measure how rapidly the filtered image changes as it crosses zero. This quantity is related to the contrast and sharpness of the intensity change.

Intensity changes at different spatial resolutions can be analyzed by varying the sizes of the two Gaussians. Figure 10 illustrates a single image and the results of filtering and zero–crossing detection that use different size DOG functions. A larger operator captures the gross structure of the image, while smaller operators capture its fine detail. This is the kind of spatial information that is accentuated by the multiple channels in the human visual system.

a.

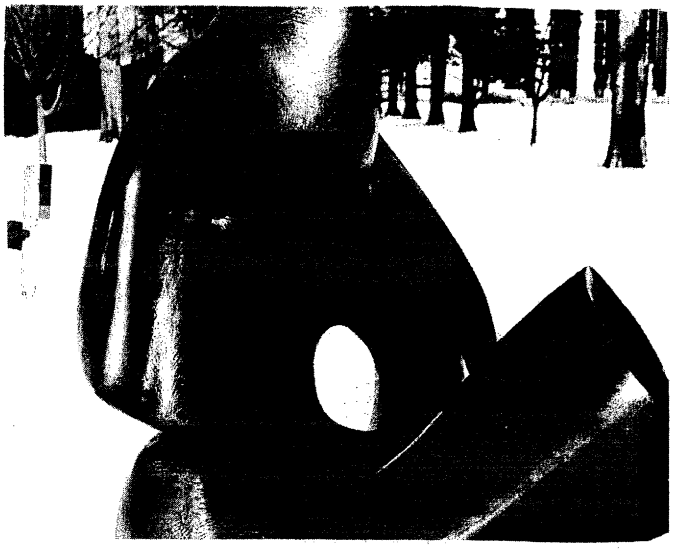


b.



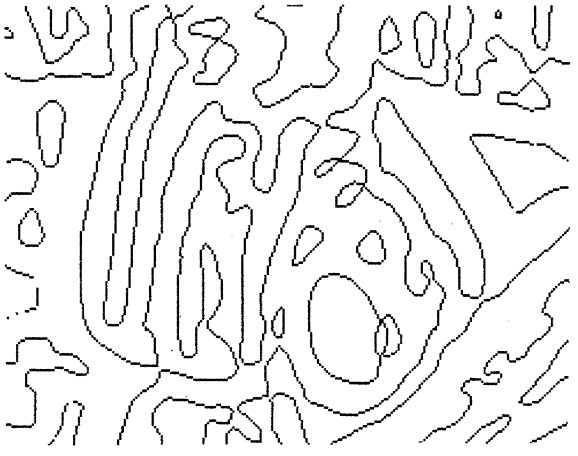
c.

Figure 9: Detecting intensity changes. (a) An image of a natural scene. (b) The result of filtering the image shown in (a) with a difference-of-Gaussians function. (c) The positions of the zero-crossings of the filtered image shown in (b).

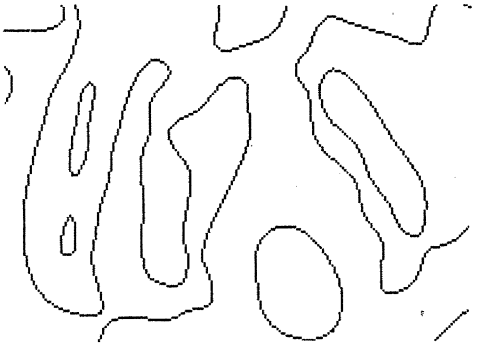


a.

b.



c.



d.

Figure 10: Using multiple operator sizes. (a) The image of a natural scene. (b), (c) and (d) The positions of the zero–crossings that result from filtering the image shown in (a) with difference–of–Gaussian functions whose central positive region has a diameter of 6, 12, and 24 image elements, respectively.

From a theoretical perspective, many operators can be used to filter the image for detecting intensity changes. The DOG function is one possibility, but in general the operator need not have a Gaussian shape and need not be circularly symmetric. Theoretical studies, however, have shown that in one dimension, an operator whose shape is given by the first or second derivative of a Gaussian[10] can be best suited for detecting intensity changes[11] (Shanmugam, Dickey and Green, 1979; Marr and Hildreth, 1980; Canny, 1983; Poggio, Voorhees and Yuille, 1985; Torre and Poggio, 1985; Yuille and Poggio, 1984a,b). There is still debate over the best operators to use in the detection of intensity changes in two dimensions. Under some criteria, the Laplacian of a Gaussian, or its DOG approximation, is best suited for the task (Marr and Hildreth, 1980; Torre and Poggio, 1985; Yuille and Poggio, 1984a,b). In other words, if the retina is filtering the image for detecting intensity changes, it is performing this function in one of the best ways possible. This observation is nontrivial. It may not shed further light on *what* operations are performed in the retina, but it does suggest *why* these operations are performed at the first stages of vision. Substantial theoretical work on edge detection presently is directed at two further questions. First, to what extent can a representation of the changes of intensity capture all of the important information in the image, and second, how can a description of the changes in the image be used to understand the physical changes taking place in the real world.

Computational studies have suggested specific models for the function of some classes of neurons in the visual pathway, which can be tested through physiological experiments. Let us consider an example of a possible model for one function of simple cells in the visual cortex. The model first assumes that the input to the visual cortex that is carried by the X system represents the spatial filtering of the retinal image with the DOG function, combined with a temporal filtering (Rodieck and Stone, 1965; Enroth-Cugell and Robson, 1966; Hochstein and Shapley, 1976; Victor and Shapley, 1979; Shapley and Victor, 1981; Richter and Ullman, 1982). The elements in this input that correspond to significant intensity changes in the image are the zero-crossings. We might therefore hypothesize that simple cells play a role in the detection of zero-crossings in the filtered image provided by the X system (Marr and Poggio, 1979; Marr and Hildreth, 1980; Marr and Ullman, 1981; Poggio, 1983).

---

[10] In one dimension the second derivative of a Gaussian can be approximated by the difference of two one-dimensional Gaussian functions.

[11] A variety of criteria have been used to evaluate the *best* operator. Some studies examine the ability of the operator to detect a step change of intensity that is embedded in a pattern of noise, where the noise might by Gaussian or uniformly distributed. Across a wide variety of different criteria, the Gaussian shape appears to be best suited for detecting intensity changes.

Specific models have been proposed that suggest how simple cells could detect zero–crossings (Marr and Hildreth, 1980; Marr and Ullman, 1981; Poggio, 1983; Richter and Ullman, 1984). A neural zero–crossing detector can be constructed straightforwardly by combining the outputs of the ON and OFF center cells. Suppose the ON and OFF center cells carry the positive and negative parts of the DOG–filtered image, respectively. A zero–crossing is a transition between positive and negative values in this fitered image. A zero–crossing is then revealed by the presence of significant activity in ON center cells adjacent to significant activity in OFF center cells. This observation led to the model illustrated in Figure 11a (Marr and Hildreth, 1980; Marr and Ullman, 1981). The outputs of adjacent ON and OFF center cells are combined through an AND operation. In this model, the cell is active only when a zero–crossing is present in the DOG–filtered image that forms the input to the cell. The ON and OFF center cells can also be arranged in columns to provide the cell with additional selectivity for the orientation of a local zero–crossing contour (Marr and Hildreth, 1980). Pharmacological and physiological studies, however, do not support the particular model shown in Figure 11a for how simple cells might combine inputs from the LGN (Sillito, 1975, 1977; Sillito *et al.*, 1980; Schiller, 1982).

The work by Sillito and his group (Sillito, 1975, 1977; Sillito *et al.*) suggests that the selectivity of simple cells for both the orientation and direction of movement of an edge or bar involves inhibitory interactions of some type. This conclusion is based on experiments showing that direction selectivity is abolished and orientation selectivity is impaired when the chemical substance bicuculline is injected into an area of the visual cortex. Bicuculline is thought to act antagonistically to the putative cortical inhibitory neurotransmitter GABA[12]. In the particular model shown in Figure 11a, orientation selectivity arises through AND–like interactions between the inputs (or an array of inputs). The model does not make explicit use of any inhibitory interactions.

The results of Schiller's (1982) experiments suggest that the sensitivity of cortical cells to the presence of edges in their receptive field arises through the interaction between cells of a single type (either ON or OFF center cells alone). This study used the observation that injection of the chemical substance APB[13] into the retina reversibly blocks the ON center cell system, thus preventing any outputs of the ON center cells (within a particular area of the visual field) from reaching the visual cortex. While the injection of APB was effectively blocking the ON center system, Schiller made the following observations of cells in the visual cortex: (1) cells that originally responded

[12] γ–aminobutyric acid

[13] DL–2–amino–4–phosphonobutyric acid

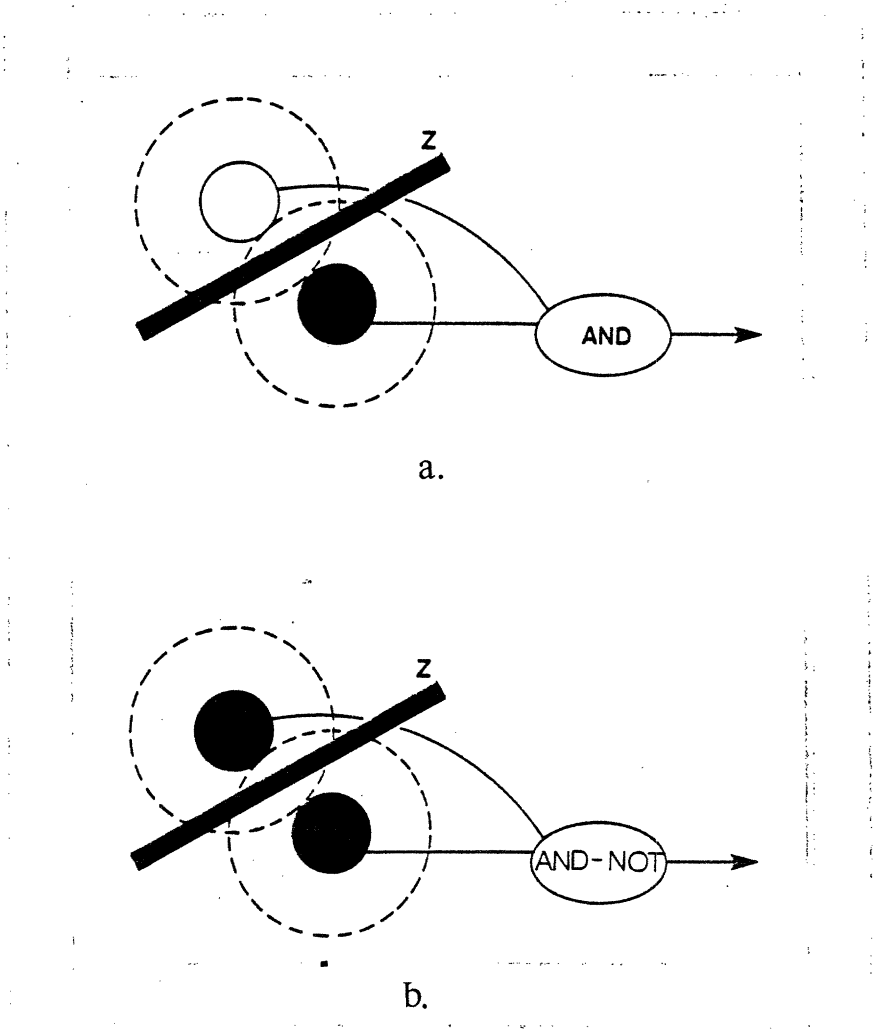


Figure 11: Simple cell models. (a) The simple cell model proposed by Marr and Hildreth, in which the responses of adjacent ON and OFF center LGN cells are combined through an AND-like operation. (b) The simple cell model proposed by Poggio, in which adjacent LGN cells of the same type are combined through an AND–NOT operation.

to edges of either contrast sign now only responded to edges of one contrast sign, and (2) cells did not loose their orientation or direction selectivity. In the model of Figure 11a, the detection of edges arises through the interaction between ON and OFF center cells. It is therefore inconsistent with Schiller's study, which suggests that simple cells can detect moving edges when only the OFF center cells are active.

The above mentioned pharmacological and physiological studies led to a subsequent model for simple cells proposed by Poggio (1983; Koch and Poggio, 1985) that combines two of the same kind of cell with an AND–NOT operation (illustrated in Figure 11b). A zero–crossing is detected when there is significant activity in, say, the ON center cells, adjacent to an area of no activity in the ON center cells[14]. The "NOT" part of the AND–NOT operation can be carried out by inhibitory interneurons, yielding a model that is consistent with the experiments by Sillito (1975 1977; Sillito *et al.*, 1980). This model is also consistent with the study of Schiller (1982), because an edge is detected through the interaction of only one cell type (either ON or OFF center cells). The model proposed by Poggio was therefore guided both by a computational analysis that showed the importance of zero–crossings, and by experimental data regarding the neural properties of simple cells.

The AND and AND–NOT operations appearing in the simple cell models of Figure 11 should not be interpreted as strict boolean logical operations, as neurons in general do not function in a discrete binary manner. The fundamental biophysical processes that underly information processing in neurons, i.e. conductance and voltage changes, are smooth functions that (with the exception of the spike) give rise to graded, analog signals. Analyzing the computations performed by neurons in terms of boolean logical operations is an oversimplified but suggestive way of representing these truly analog operations (Koch and Poggio, 1984, 1985).

The simple cell models described above have stimulated physiological experiments to test the underlying zero–crossing hypothesis more carefully (Richter and Ullman, 1984). The experiments relied on the fact that zero–crossings in a DOG–filtered image do not always correspond to edges in the original image. Due to the smoothing of nearby edges, spurious zero–crossings sometimes occur where no real edge exists in the image. If some simple cells detect zero–crossings, they should respond to these spurious zero–crossings. The stimulus used by Richter and Ullman (1984) is a gray–level "staircase" composed of two adjacent step changes of intensity, as shown in Figure 12a. The one–dimensional intensity profiles (cross–sections of the actual stimuli used) are shown in Figure 12b for

[14]The model proposed by Poggio also includes a mechanism for the selectivity of simple cells for the direction of motion of a stimulus (Koch and Poggio, 1985).

a range of separations between the two edges. The results of filtering these stimuli with the DOG function shown in Figure 12c[15] are shown in Figure 12d. When the separation between the two edges is small (rows 1 and 2 of Figure 12d), they cannot be resolved by this DOG filter. That is, they give rise to a single zero–crossing, indicating the presence of only a single edge in the stimulus. When the separation between the two edges is large compared with the size of the DOG filter (row 6 of Figure 12d), they are analyzed almost independently. Two distinct responses to the two edges appear in the filtered profile — it decreases gradually through zero between the locations of the two edges, without giving rise to a significant zero–crossing. At intermediate separations (row 4 of Figure 12d), three distinct zero–crossings appear in the filtered profile. Two are associated with the actual intensity steps and a third of opposite sign is located at the middle of the plateau between the two. This "extra" zero–crossing indicates the presence of a change of intensity (or edge) that does not exist in the original intensity profile.

The double–edge stimulus of Figure 12 was used to test the hypothesis that some simple cells detect zero–crossings. Suppose that a simple cell responds only when a vertically oriented edge that is dark on the left and light on the right is moved from left to right across the cell's receptive field. If the cell detects zero–crossings, it should respond whenever adjacent negative and positive areas appear in the DOG–filtered image (with the negative area on the left). Suppose that the staircase stimulus of Figure 12a is moved in the preferred direction. The zero–crossing hypothesis predicts that when the two intensity edges are close together, the cell should respond only once to the two–step stimulus. When the edges are sufficiently separated, the cell should respond to each of the two edges, giving two distinct responses for each single sweep of the stimulus across the cell's receptive field. Suppose that the sign of contrast of the stimulus is then inverted, as shown (in one dimension) in Figure 13a. The result of filtering this inverted stimulus, for an intermediate separation between the two step changes of intensity, is shown in Figure 13b. For the inverted stimulus, the zero–crossing hypothesis predicts that if the two edges are close or sufficiently separated, the cell should not respond at all, because a zero–crossing of the appropriate sign of contrast never appears in the cell's receptive field. For intermediate separations, however, there appears a zero–crossing of the appropriate sign, to which the cell should respond, even though no edge of the appropriate sign of contrast exists in the original stimulus (see Figure 13b).

Richter and Ullman (1984) tested the zero–crossing hypothesis for a subclass of sim-

[15] A slightly asymmetric DOG function was used, which incorporates a temporal delay between the responses of the center and surround Gaussians (Richter and Ullman, 1982).

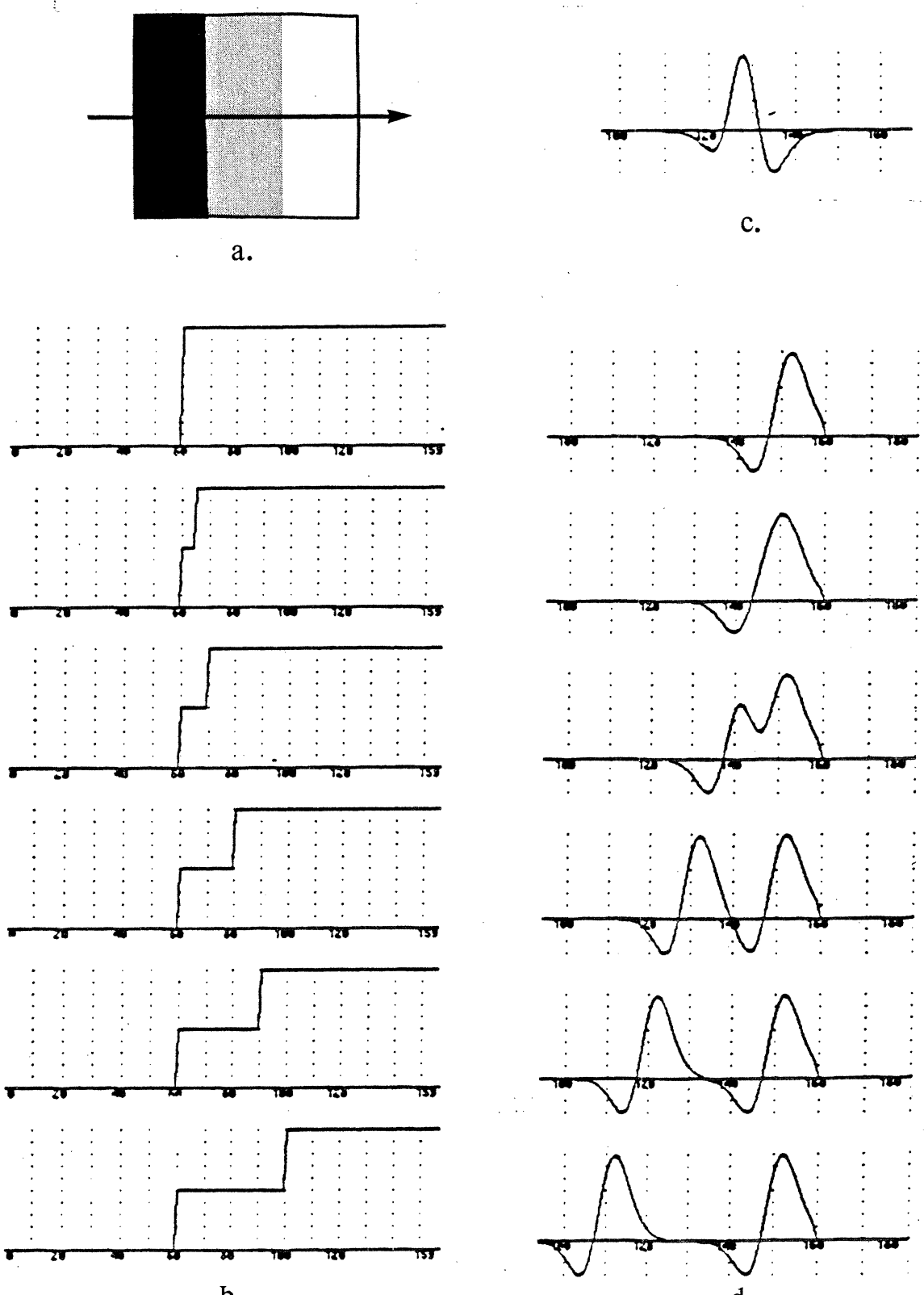


Figure 12: (a) The "staircase" stimulus used by Richter and Ullman, consisting of adjacent bars of different intensities. (b) The graphs represent the cross-section of the intensity distribution across the bar pattern, for a range of separations between the two step changes of intensity. (c) An asymmetric difference-of-Gaussians function. (d) The results of filtering the patterns shown in (b) through the difference-of-Gaussians function shown in (c).

Figure 13: Testing the zero-crossing hypothesis. (a) The staircase stimulus of Figure 12, with its contrast inverted so that the step changes of intensity are light on the left and dark on the right. (b) The result of filtering the profile in (a) with a difference-of-Gaussians function of intermediate size.

ple cells that were "edge-specific" in that they respond preferentially to edges of light of a particular orientation and sign of contrast. In the classification introduced by Schiller *et al.* (1976), this subclass includes the simple cell type S_1, which is also selective for the direction of motion of the edge, and type S_3, which responds to edges moving in both directions. In electrophysiological recordings from edge-specific cells in the cat, roughly half of the cells (28 out of 55 recorded) showed a clear response to the extra zero-crossing present in the staircase stimulus with its contrast inverted. This result suggests that the zero-crossing hypothesis may be plausible for some simple cells. The experiment by Richter and Ullman does not yet rule out alternative hypotheses about the function of simple cells. Other models suggested by Spitzer and Hochstein (1985) and Movshon (personal communication), for example, may also account for these experimental results. Although an ideal experiment would discriminate between alternative hypotheses, this example illustrates how a computational theory can lead to a specific model of the function of neurons in the visual pathway and can provide testable predictions for physiological experimentation. A recent study by Hochstein and Spitzer (1984) also provides experimental evidence regarding the possible role of simple cells in the analysis of zero-crossings. As a result of these experiments, Hochstein and Spitzer proposed that simple cells may behave as zero-crossing *filters*, in that they respond strongly in the presence of zero-crossings, but also respond weakly in the presence of other features in the input from the LGN.

Models for zero-crossing detection such as the one described above also have stimulated psychophysical work, aimed at showing what visual information is extracted from the retinal image. For example, Watt and Morgan (1983a) studied the way in which

spatial position is assigned to features in the retinal image. They considered several theoretical models for spatial localization and designed stimuli composed of bars of different luminances that would discriminate among the theoretical models. The performance of human observers in these experiments is consistent with a model that encodes the occurrence and location of zero-crossings in the second derivative of the retinal image. Later studies of the ability of human observers to judge other spatial properties of intensity variations in the image, such as the extent of blur of the intensity changes, suggested that peaks in the second derivative of the retinal image also may be used to encode retinal image information (Watt and Morgan, 1983b). An experiment by Mayhew and Frisby (1981) also suggested that peaks in the second derivative of intensity may be used in the analysis of binocular stereopsis. Recent experiments by Morgan *et al.* (1984) and van Santen and Sperling (1984) addressed the question of what spatial features the visual system uses to measure motion. In all of these psychophysical studies, theoretical models of the extraction of image features provided critical input to the design and interpretation of the experiments.

To summarize, early progress on the problem of detecting intensity changes was made independently in computer vision, perceptual psychology and visual neurophysiology, but much greater progress has come since the observations of these three fields were brought together. Insights about early visual processing in biological systems have led to more general and reliable methods for edge detection in computer vision systems. At the same time, the computational analysis of the early stages of visual processing has led to productive psychophysical and physiological experiments perhaps offering a better understanding of the function of neural mechanisms in the visual pathway. Investigators may alternate many times between theoretical models and experiments before finding models that are consistent with all experimental data. If the models are stimulated by what is needed in the system from a computational perspective, then when feasible physiological examples are found, we will have a deep understanding of why particular mechanisms exist and the role they play in visual function.

3 The Study of Motor Control

Although vision and motor control represent different areas of research in terms of theories, mechanisms, and experimental procedures, the underlying approach to their study by the computational paradigm is the same. A primary emphasis is placed on discovering and examining all possible natural constraints, and then tracing the implications of these constraints on control mechanisms. Insight into the motor control problem is obtained by the development of competence theories, by way of computer simulation and then implementation on actual mechanical hardware.

The computational approach to motor control is strongly interdependent with robotics. Both fields share the goal of the intelligent translation of perception into action. Robotics provides a convenient laboratory for developing and testing control principles. Although the differences in mechanical structure and computational architecture between biological systems and machines might at first seem to differentiate robotics from motor control, at a certain level of abstraction the problems encountered are the same. Motor control is first and foremost a mechanical problem. The body is composed of linked segments with attributes of mass and geometry, which accelerate in a gravitational field and interact with objects in the environment. Just as in robotics the biological motor control system must have developed to reflect these mechanical constraints, even though in this instance control signals are sent to muscles rather than to motors.

The above considerations indicate how robotics has contributed towards the understanding of motor control at a higher level than muscles and nerves, in line with the "top-down" nature of the computational approach. External constraints on movement necessarily have been defined, since movements frequently contact environmental surfaces. Mechanical constraints of linkage geometry and mass also have been examined in great detail. Even at the level of actuator constraints, functionally equivalent models of motors and muscles have sometimes been proposed. Recent trends in the design of tendon-actuated robot hands are actually bringing the respective mechanical structures closer to biological counterparts (Jacobsen et al., 1984).

Properties of the neuromuscular system impose intrinsic limits on what models may be proposed, so that concepts developed in robotics must be carefully evaluated for purposes other than general background since some may be biologically inapplicable. At the same time, biology offers important clues to investigators of robotics, since human motor performance generally far outstrips robot performance. In the sections that follow, several sources of constraints that operate at the neuromuscular, mechanical, and external levels are examined. The section goes on to present as a competence model a hierarchical movement planning and control structure adapted from robotics, so as to

give an example of how these constraints can be accomodated. This structure is examined with respect to its implications for biological motor control and its accomodation with experimental results.

3.1 Features of Motor Control Research

The understanding of biological motor control has proven stubbornly difficult, and we cannot yet establish a direct analogy between neural processing and computational studies. Whereas in vision it is known that processes of edge extraction, stereopsis, and optical flow exist, in motor control no consensus agrees upon the fundamental transformations. It is an open question whether biological processes for inverse kinematics[16] or inverse dynamics[17] exist, or whether the nervous system plans movement trajectories explicitly on a point-to-point basis. Indeed, we do not know even whether control operates on variables at the level of muscle, joint, or endpoint, and for a given level whether these variables specify stiffness, length, velocity, force or torque (Stein, 1982). For motion-related sensors and neuronal centers, basic questions such as the influence of muscle spindles in movement remain unresolved (Hulliger, 1984). The specific contributions of motor cortex, cerebellum, basal ganglia, and spinal cord in motor computations are even less well understood.

A significant difference between motor control and vision is that the former is not just a pure information processing problem. Between motor performance and neural processing lies a set of complex biomechanics that greatly enhances the difficulty of relating motor events directly to neural events. Properties of this biomechanics are integral to the formulation of a motor plan. Said another way, motor control cannot be understood without knowing the biomechanical properties of the system and how these properties influence and are accommodated by the motor control system.

When a muscular movement is observed, it is not necessarily the reflection of a neural process. A simple analogy would be a spring, which oscillates with an attached mass purely due to mechanical properties. It has been proposed that muscles possess spring-like properties that can be organized to realize complex movement goals with simple forms of control (Hogan, 1982). In a humorous vein McMahon (1984) has suggested that the function of neural control during running is to prevent disruption of the natural mechanical resonances of the system. Motion at one joint is also influenced by motions at other joints, due to the effects of complex dynamic interactions (Hollerbach and

[16] The transformation from endpoint variables to the corresponding joint angle variables.

[17] The transformation from joint positions, velocities, and accelerations to joint torques.

Flash, 1982). The elbow will flex passively, for example, in reaction to acceleration at the shoulder, and vice versa.

In addition to their differences in output functions, movement is voluntary and discontinuous, while vision is involuntary and continuous (at least at the lower levels). A paralyzed and anesthetized animal can visually process a pattern without any act of volition, and this continuous and repeatable input may be traced through the neuronal circuitry to infer its associated transformations. Movement, by contrast, once executed is finished, hence is discontinuous; subsequent repetitions may differ and give rise to varied neuronal activity. Since movement is primarily a voluntary activity, alterations of the CNS by drugs or surgery severely compromise the ability of the system to perform naturally. It remains a major controversy whether movement features following a neuronal lesion indicate the role of the lesioned center in motor control, or represent a totally different strategy of the animal compensating as best it can with the remaining circuitry.

The psychophysics of movement is less well developed than that of vision, primarily because natural movements are difficult to measure. Movements often must be reduced to the simplest cases, usually about single joints, merely because of the difficulty in recording kinematic features and of applying perturbations except in simple configurations. EMG signals are hard to interpret, especially during active movement. Nevertheless danger exists in simplicity: limiting studies to single-joint movements may lead to too narrow a view of what is involved in motor control.

Fortunately, experimental techniques have improved substantially in recent years and should ameliorate many of the past limitations. Movement monitoring systems such as the Selspot system[18] (Atkeson and Hollerbach, 1985) allow measurement of kinematic features of unconstrained, natural movements. Neuronal recording techniques have improved; floating electrodes for example allow spinal recording during natural cat locomotion (Hoffer et al., 1981). One remaining difficulty, however, is the application of perturbations to ongoing movement, since almost by definition a natural movement cannot be constrained by an apparatus that is to apply the perturbations. Though perturbations are currently limited to one or two-joint movements for experimental study, nevertheless there has been much new information about biomechanical properties and reflexes by the use of sophisticated engineering analysis (Kearney and Hunter, 1983).

A remaining general consideration is that there are many sensorimotor systems and their interrelationships are often unclear. Eye movement for example may have little

[18] An optoelectronic stereo camera system, produced by Selcom of Sweden, that senses infrared led markers attached to a limb by means of lateral-effect diodes in each camera.

in common with arm movement, which in turn may differ substantially from locomotion. The eye is a comparatively simple and predictable mechanical object, as its mass never changes and its movement is confined to orbital rotation. By contrast, the arm is kinematically complex and varies in its load conditions due to gravity, grasped objects, and environmental contact. In locomotion it is not clear that the leg trajectory need be controlled similar to the arm trajectory. So far, several different theories with little in common have been proposed for these various sensorimotor modalities: linear control theory for eye movement (Robinson, 1973), potential field models for arm control (Hogan, 1982), and oscillation models for locomotion (Grillner, 1975). Ultimately one hopes to find unifying principles that underly all of motor control, but such rules can emerge only after a more thorough understanding of the individual systems.

The remainder of this section focuses primarily upon control of human arm movements. Other sensorimotor modalities that could have been discussed in terms of the computational approach are locomotion (Raibert, 1984) and hand control (Jacobsen et al., 1984, 1985; Salisbury and Craig, 1982). Research into one and four legged hopping machines is generating new ideas about modular processes in locomotion, while research into the design and control of four-fingered, tendon-driven robot hands is providing information about elementary hand functions and the use of contact sensing.

3.2 Natural Constraints in Motor Control

Natural constraints confront the motor control system at several levels: neuromuscular, mechanical, and task. The neuromuscular level reflects the mechanical and computational properties of the biological system. The mechanical level views limbs as mechanical linkages and analyzes them from a standpoint of kinematics and dynamics. The task level focuses on how endpoint positions and forces should evolve in response to environmental goals and constraints. When comparing control of movement between biological and robotic systems, the considerations are similar at the mechanical and task levels, but they differ at the neuromuscular level since the computational and actuational characteristics of robots are so dissimilar.

All sources of constraints must be identified for their possible effects on the nature of the biological motor controller. The following sections give examples of constraints at each level, and indicate how they influence hypotheses about motor control.

3.2.1 Neuromuscular Constraints

The motor control system must sense its own machinery and be able to forge a solution within its limitations. The mechanical machinery includes muscles, joints, and limbs. Muscle is a complex tissue, whose contraction depends on force, velocity, and level of activation. The contribution of passive tissues such as ligaments and tendons must be considered. Individual muscle fibers display a variety of architectures when assembled into a whole muscle, such as pinnation and compartmentalization (Loeb, 1984). Redundant musculature surrounds most joints, e.g. the shoulder joint contains 18 muscles. Over half of all muscles pass over two or more joints. Some muscles have elaborate three-dimensional trajectories during contraction; the normal biomechanical assumption of straight-line trajectory between origin and insertion would predict the wrong direction of torque production (Wood, Meek, and Jacobsen, 1984). Further, most joint articulations do not satisfy ideal models such as a hinge joint (knee) or a spherical joint (wrist); the clavicle moves with five degrees of freedom, which is close to that of a free body.

Signal transmission and processing delays in the nervous system have far-reaching implications on how the motor system can conduct real-time control. While spinal feedback loops for arm movement have a latency of 25 msec, these feedback loops turn out to have too low a gain[19] to operate effectively to counteract movement perturbations (Bizzi et al., 1978). The more substantial long-latency responses of 80-100 msec are too long to serve effectively as closed-loop feedback, because control under conditions of substantial feedback delays would be unstable (Hollerbach, 1982). For moderately fast arm movements, by the time a corrective response can act, the limb will have reached a new state for which the response is inappropriate. Although delays can be compensated if higher-order derivatives of the error are known, it is unlikely that the nervous system could accurately compute these derivatives (Arbib and Amari, 1985).

In the face of the above arguments and a variety of experimental evidence, it has been concluded that fast to moderately fast arm movements must be controlled open-loop[20]. Feedback would not serve for fine tuning of ongoing movement in the classic servo sense, but could monitor the movement for global adaptation (such as estimating a load mass), local adaptation (such as refinement of points along a repeated trajectory), or reprogramming after major disturbances. These limits on feedback efficiency make inappropriate attempts to apply linear control theory where instantaneous and accurate

[19] Sanes (1983) indicates that the short-latency spinal loop may be more effective during small perturbations or incremental movements.

[20] The term *open-loop* refers not to the absence of feedback but to the role feedback plays.

feedback is typically assumed to biological motor control. The only alternative seems to be that the motor control system has constructed a system that allows accurate predictive control.

Interestingly this biological solution goes against developments in modern control theory, where it is often argued that such a complex system can only be handled by robust control. A robust controller relies on feedback because thus far a sufficiently accurate and useful model system has not been constructed (Slotine, 1985). It appears that the biological solution may provide an alternative viewpoint on how a problem can be solved and may prevent us from believing too strongly that our artificial constructs are the only way of proceeding.

Given the complexity of the biological machinery, some have characterized biological motor control as a smart controller for sloppy hardware. According to this view, the controller must fight with an unpredictable and uncooperative system to achieve a successful movement. Whatever the properties of the system might be, the task would be to refine the controller sufficiently to overcome the system's natural tendencies.

It is unlikely that this view is either correct or workable. The motor control machinery is anything but sloppy, and the more we learn about muscles, tendons, and sensors, the more we realize advantages they have over man-made hardware. Rather than considering system properties as making control more complex, perhaps these properties are actually adapted to accomplishing a motor task and indicate something fundamental about the task (Jacobsen et al., 1985). The way tendons in the fingers split and individually route over bumps at the joints may reflect a useful geometrical computation, for example, the ratio of joint movements. Furthermore, the particular combination of active and passive stiffness in the muscle-tendon system may allow stable recovery from unexpected collisions. Natural selection may have generated biomechanics of leg muscles for optimal locomotor efficiency (Loeb, 1984). It is a maxim in mechanical engineering that design must interact with control, and the biological system may have evolved this maxim to the furthest degree.

3.2.2 Mechanical Constraints

Above the level of the muscles and nerves, the body can be considered as an assembly of mechanical links and joints. Movement of these links must satisfy the geometrical constraints of the environment and the goals of movement. Nevertheless, these links are inertial objects, with attributes of mass, center of mass, and inertia. These links are acted upon by gravity, and as mentioned earlier their dynamic interactions complicate joint torque production. Just as with neuromuscular constraints, the mechanical

constraints of kinematics and dynamics restrict the range of possible control strategies.

The whole problem of kinematics is the nonlinear transformation between end positions, orientations and joint angles. The difficulty of this tranformation has far-reaching implications for motor control. An object in space can be located by six variables, three for position (such as Cartesian x, y, z coordinates) and three for orientation (such as roll, pitch, and yaw angles). To grasp such an object, a linkage system also must have at least six degrees of freedom. The *inverse kinematics* problem is, given the location of an object in space, to find the joint angles that correspond to the arm at that location. When the linkage has more than six degrees of freedom, it is *kinematically redundant* because there are more degrees of freedom than absolutely necessary for general positioning. Redundancies are useful to avoid obstacles, eliminate internal singularities, and avoid joint limits.

The inverse kinematic transformation is only computationally efficient if the mechanical linkage contains certain kinematic arrangements (Pieper, 1968). One such arrangement is a spherical joint, usually at the wrist, which allows separation of positioning from orienting. It is probably no accident that humans have spherical wrist joints to accomodate a roll motion in the forearm, a pitch motion at the wrist (flexion-extension), followed by a wrist yaw motion (abduction-adduction). The human arm actually possesses redundant motion because it has seven degrees of freedom (not counting body movement): three degrees at the shoulder joint, a single degree at the elbow joint, and the spherical wrist. It has been argued that this particular kinematic arrangement is optimal in terms of the advantages of redundancies mentioned above (Hollerbach, 1985). At the same time that redundancies bring advantages, however, they make calculation of the inverse kinematics transformation more complicated because of the necessity of resolving the redundancy (Hollerbach and Suh, 1985).

The first level of abstraction above muscle activation is dynamics, which relates torque production at the joints to desired joint position, velocity, and acceleration. What makes dynamics complex is the presence of interaction torques, due to inertial, centripetal, and coriolis forces. Inertial forces are the normal actions and reactions that result whenever a body is accelerated, but for a multi-joint linkage, acceleration at one joint creates a reaction torque at other joints. Centripetal forces are proportional to squared velocity and are analogous to inwardly directed accelerations; an example keeps a ball whirled around on a string in a circular orbit. The forearm also represents a body kept in orbit about the shoulder joint, attached by the upper arm, and must leave centripetal torques acting at all joints. Coriolis forces arise whenever two rotating systems interact, for example, the rotation of the earth with north-south movement

of hot and cold air, which gives rise in the northern hemisphere to counterclockwise vortical forces. The rotating sytems in the arm are the upper arm, forearm, and hand, which interact to yield a complex combination of coriolis forces.

The interaction of these several forces is usually overlooked in motor control, partly because these forces are complex. Also, investigators might be hoping that the interacting forces are ordinarily insignificant or can be overcome with feedback. But as Hollerbach and Flash (1982) showed, all three types of interaction forces operate during ordinary movement and cannot be ignored. Moreover, they pointed out that linearization of dynamics cannot be justified on the basis of movement speed, because they showed the dynamic interactions to be speed invariant. This contradicts the normal assumptions in the robot control literature, where investigators have attempted to fit arm dynamics into linear control theory more because the latter is a well-developed area than because it is well suited.

The motor control system cannot treat dynamic interactions as perturbations or errors to be corrected by feedback, because of the transmission delays mentioned earlier. Even if feedback were faster, it is unlikely that a controller could ignore dynamics without running into stability problems. It might be justifiable to ignore link dynamics if muscle dynamics were dominant. Whether one can do so depends on the particular circumstance. The fingers are relatively light compared to the muscles that activate them; combined with frictional losses of the tendons routing all the way from the fingers to the forearm, it is likely that muscle/tendon dynamics dominate the finger dynamics. Said another way, finger muscles overpower finger mass. For the arm or leg, however, the limb masses are substantial and lead to significant link dynamics.

It is fair to say that the mechanical constraints in motor control have been underemphasized relative to the neuromuscular constraints. Part of the reason is the restriction to studies of single-joint movement, where issues of kinematics and dynamics are trivial. In such studies all levels of analysis are the same: force is directly proportional to acceleration, and there are no dynamic interaction forces. Trajectory planning degenerates to control of one position variable. On the other hand, for multi-joint movement force or torque and acceleration become no longer proportional. Trajectory planning becomes a complex problem of relating joint angles to externally defined positions. Even planar two or three joint movement is considerably simplified over three-dimensional movement, because orientation is so much more complex in three dimensions than in two.

3.2.3 External Constraints

Movement is not just a matter of freely generating trajectories, but is constrained both by task demands such as accurate throwing and by physical contact with external surfaces. When an object of unknown weight is picked up, the dynamic characteristics of the arm suddenly change. The motor control system must quickly estimate the inertial parameters of the object relative to the hand's grasp and update its internal model to achieve a skillful movement. Similarly, when a pointer is grasped, the kinematic parameters of the arm are changed suddenly. Again, these parameters must be estimated quickly and incorporated to modify the kinematic solution relating endpoint and joint angles.

The geometry of the external world constrains how movement may take place, by defining a set of natural coordinates by which to plan the action. Picking up a cup requires definition of the cup position and orientation, to be matched by an approach direction and a grasp. The cup must be kept level to avoid spilling when transported, and the motion should be fairly straight to minimize angular accelerations that could also lead to spilling. As the cup or other object is moved, obstacles must be avoided and a path found through a cluttered environment. Often in robotics straight-line Cartesian paths are preferred because it is easier to predict the consequences of movement in terms of avoiding obstacles. Real-time constraints must be matched as well, as in catching a flying object where the hand must achieve a specific position at a specific instance of time.

Motion can also be constrained by prolonged contact with an environmental surface. Writing on a blackboard requires that movement can take place only parallel to the board and not into or away from it; the board sets up a natural coordinate system defining allowable directions of movement plus another direction in which force can be generated, as discussed below. We can match our movements and force application to the external coordinates of a light socket so as to screw in a lightbulb in practically any position and orientation – above our heads, to the side, or when we're upside down. Opening a door requires that the hand follow the natural circular trajectory of the door handle.

If the generation of endpoint positions were the only problem in motor control, it would be hard enough, but it is only half the problem. The other half is the requirement to generate endpoint forces, arising precisely when environmental surfaces are contacted. For blackboard writing, the environment signals that it is forbidden to move through the board by pushing back as hard as the hand pushes into it. During contact the normal six degrees of positioning freedom are reduced, yet it is a fundamental law of mechanics that

the lost freedoms of position are converted into freedoms of force and torque. Writing on the board is an example of point contact with five positioning freedoms, by the hand translating in two directions and rotating the chalk point in three directions. Against the board one can generate a normal force but not a displacement; this single force freedom plus the five position freedoms add to six by necessity.

Force control or compliance is recognized as a fundamental issue in robotics (Mason, 1982), and in some sense is considered as a somewhat different problem from the generation of unconstrained trajectories. Reliance must be placed on contact sensors instead of on position sensors, and the servo response to contact must be much more rapid than to position errors because very high contact forces can arise in little time with essentially no displacement. A mechanism that could serve as an alternative or complement to an active force servo would be passive compliance by actuators, transmission elements, or structures. Such an approach seems particularly pertinent for biological motor control. The difficulty with passive compliance schemes, however, is to arrange the compliant elements in a useful and flexible way, not just for one kind of contact condition, but for many contact conditions. Later, it will be considered how the spring-like properties of muscle ensembles could be organized in this way.

A general form of motor control is hybrid force/position control, where certain degrees of freedom are controlled for position and other degrees of freedom are controlled for force (Raibert and Craig, 1982). The task in environmentally constrained motion is to generate a movement plan with the best available geometrical information about external surfaces, but to recognize that one's external model will be uncertain and that one will have to comply with contact conditions when the model and the actual external geometry differ.

By and large, investigators in biological motor control have emphasized control of position rather than control of force during constrained movement. Yet as mentioned earlier, most experimental movements are restricted to one or a few freedoms of position because of measurement difficulties. Thus since the excluded position freedoms are merely transformed into force freedoms many workers have inadvertently been studying compliant motions without recognizing that such was the case. The arm generates a six-dimensional force/torque vector at the hand, but only the vectorial component projected onto the instantaneous motion axis would be observed in control of position studies. An important question to be resolved in biological motor control is whether, as in robotics, different considerations govern force control and position control, or whether some common principle underlies both (Hogan, 1984). The danger is that by not focusing on freedoms of force in an experimentally constrained movement, one

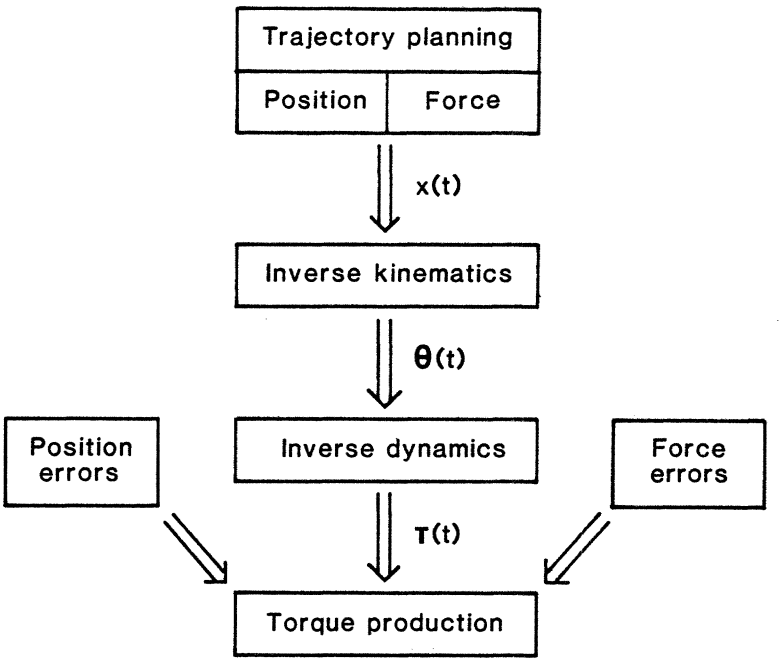


Figure 14: A modular planning and control structure for robot arm movement. A trajectory is planned in hand coordinates, synthesizing a hybrid force-position strategy. The endpoint trajectory $x(t)$ is transformed into a joint trajectory $\theta(t)$ by solving the inverse kinematics. The feedforward torques $T(t)$ are then found by solving the inverse dynamics, and are corrected by feedback for force and position errors.

misses an essential component of the motion.

3.3 Movement Planning Hierarchy

The sources of natural constraints fall roughly into a hierarchy, and in robotics a movement planning and control structure has evolved that directly reflects these different levels. The control structure consists of an object level, a joint level, and an actuator level, and is represented as a sequence of transformations in Figure 14. Trajectory planning takes place at the object level. External task constraints are synthesized by planning a time sequence of endpoint positions and forces that form a correct interface to the geometry of the external world.

At the next level, the joint level, the time sequence of endpoint coordinates are transformed into a time sequence of joint angles by solving the inverse kinematics problem. At the actuator level, the time sequence of joint angles is converted into a time sequence of joint torques by solving the inverse dynamics problem and by feedback correction of errors based on position and force. Mechanical constraints are therefore synthesized at both the joint and actuator levels. Another aspect of computation at the actuator

level is to convert the joint torques to commands appropriate for each actuator and its ultimate controller, such as current for an electromagnetic motor. How a joint torque is transformed into a motor torque depends greatly upon the particular actuation system, and would correspond to the neuromuscular constraints in biological motor control.

The purpose of elaborating this hierarchy of movement planning and control derived from robotics is to provide a competence model that applies to general motion control. The elements of this model define conceptual stages in processing, which serve at the very least as descriptions of the motor task if not as prescriptions for a control strategy. This framework then allows one to consider whether the biological motor control system can exhibit the same level of flexibility in motion control, and if not how any limitations may be reflected in shortcuts or specific solutions to elements of this structure.

3.3.1 Object Level

At the highest level, motion of an endpoint or a grasped object alone is planned, without specific consideration that an arm is required to move the object or endpoint. It is as if Adam Smith's invisible hand were applied, not to the economy, but to placing an object in a desired position. The object level has also been called the ideal effector level, because it is presumed that the effector can generate whatever forces or positions are required by the task.

Planning at the object level therefore proceeds by analyzing the natural constraints of the task. A geometric analysis should indicate what are the positioning freedoms and what are the force freedoms, as the first step towards synthesizing a hybrid force-position control. For example, suppose it is desired to slide an object along a surface. A *generalized spring* strategy is one way to synthesize forces normal to the surface while generating positions tangent to the surface. In response to normal displacements caused by movement or modeling errors, the object moves like a spring and generates proportional forces. It is called a generalized spring strategy (Mason, 1982) because action of the spring can be placed in any arbitrary direction, according to the contact conditions. Moreover, contact forces and resulting displacements can be combined in a general manner; for example, the object could have been made to rotate in response to a contact force in order to place an object flush on a surface.

Specification of motion is best done by employing external variables such as Cartesian position and orientation, since these are most conveniently applied to capture the geometrical constraints of the environment. Often trajectories are made to take on simple features such as straight-line motion in these variables, because then consequences of movement such as avoidance of collisions can often be predicted.

Of course, motion of the endpoint cannot be planned in complete isolation of characteristics of the actuation and mechanical linkage, and some of the hardest problems in robotics involve propagation of constraints from a lower to a higher level. Finding the minimum-time path, for example, depends critically on properties of the actuators (Sahar and Hollerbach, 1985; Rajan, 1985). Collision of the arm with obstacles must be considered as well as collision of the endpoint (Lozano-Perez, 1982). Nevertheless, attributes of the lower levels are largely considered to provide the broad boundaries of movement, such as maximum reach, payload, and acceleration, and as long as movement stays safely within these boundaries they do not overly limit the planning process.

3.3.2 Joint Level

Once time sequences of hand positions and orientations have been specified, they are transformed into a corresponding time sequence of joint angles, through solution of the inverse kinematics problem. For the generalized spring strategy, a time-varying joint stiffness can be found that realizes the desired hand stiffness. Several problems complicate inverse kinematics: singularities, redundancies, joint limits, and obstacles.

Singularities are manipulator configurations for which there are fewer than six degrees of freedom. They arise when joint axes are aligned in such a way that a particular direction of motion becomes impossible. Workspace boundaries are always singular; for example, if the elbow is straight then no further radial motion of the wrist is possible. A larger problem consists of singularities in the interior of the workspace; for a typical rotary-joint manipulator, they occur when the wrist is straight or the wrist point is over the shoulder. With such singularities the inverse kinematic velocities cannot be solved, and hence singular points cannot be utilized in a trajectory and must be avoided. Large portions of the rotary manipulator's workspace become useless in this way.

The best solution to singularities is to add extra degrees of freedom and hence to make the manipulator redundant. Seven is the smallest number of degrees of freedom that eliminates all interior singularities. As mentioned earlier, the human arm has seven degrees of freedom (not counting shoulder movement), consisting of spherical wrist and shoulder joints and a rotary elbow joint. The extra degree permits a self-motion, which is an internal linkage movement that does not move the endpoint. With the hand fixed, the elbow point can move in a circular arc about a line joining the shoulder point to the wrist point. This allows interior singularities to be eliminated, because if a singular configuration happens to arise then a self-motion can be exercised to move the arm out of the singular configuration. Hence all interior workspace points can be utilized.

Redundancies are also useful for avoiding joint limits and obstacles. The self-motion

can be used to find a new set of joint angles in the event that one of the angles approaches its limit, again without moving the endpoint. The forearm and upper arm define a major plane of movement; the self-motion rotates this plane about the shoulder-wrist line. If an obstacle lies in the major motion plane, then it might be possible to avoid it by rotating this plane (Hollerbach, 1985).

In terms of the calculations involved in redundancy resolution, the main focus has been on the generalized inverse technique (Liegeois, 1977). This is an instantaneous optimization method, and has been used to avoid joint limits, to partition endpoint variables into high and low priority, to avoid obstacles, to minimize kinetic energy, and to minimize torque production. Nevertheless, because the generalized inverse optimizes a local trajectory point, it is possible that the solution will not remain globally optimal across a whole trajectory. In fact, Hollerbach and Suh (1985) showed that during torque optimization a whipping action develops gradually that thrusts the endpoint off the intended path.

3.3.3 Actuator Level

From a time sequence of joint angles, the corresponding time sequence of joint torques is found by solving the inverse dynamics problem. Initially it was thought in robotics that the dynamic equations are too complex to solve in real time, but it is now known that highly efficient recursive formulations exist that are of linear complexity in the number of degrees of freedom (Brady et al., 1982; Hollerbach, 1980). Furthermore, if the kinematic configuration is simple and the mass distributions are symmetric, as is true of the human arm, the dynamic equations become drastically simplified down to a manageable number of operations (Hollerbach and Sahar, 1983). Even if this number were not adequate, it is possible to recast the dynamic equations into a parallel architecture executable in time proportional to one multiplication and 3 additions, after an initial startup time (Lathrop, 1985).[21] Thus there is no longer any question in robotics about computing dynamics in real time.

With solution of the above problem, research in robotics has shifted instead to questions of whether a sufficiently accurate dynamic model of the robot can be formulated to be useful for control. If the model of the robot is not sufficiently accurate, then a predictive control based on this model will lead to substantial errors and instabilities.

(1) The dynamic equations for a robot arm require knowledge of the inertial parameters for each link.

[21] If feedback errors are processed through the dynamics, then the startup time becomes the critical factor.

It is seldom known what the inertial parameters of robot links are, since manufacturers typically specify only the kinematic parameters and the inertial parameters are incidental attributes of design. Fortunately automatic calibration methods have recently been developed that infer these parameters as a result of movement. Since the inertial parameters appear linearly in the dynamic equations, they can be estimated by least squares by relating joint torques or forces to joint velocities and accelerations (An, Atkeson, and Hollerbach, 1985; Olsen and Bekey, 1985). A related problem is load estimation of objects picked up by the manipulator, since a change in the load changes the kinematic and dynamic characteristics of the manipulator. Although the inertial parameters of loads can be derived through joint torque sensing as above, it is more accurate to use full wrist force-torque sensing (Atkeson, An, and Hollerbach, 1985; Mukerjee and Ballard, 1985).

The above methods can be implemented on-line or off-line,[22] and require no special calibration movements. The accuracy of the estimation depends on how well joint torques or forces and accelerations can be sensed and on how fast the robot can accelerate. Yet inaccuracies in inertias may not pose a problem for control, because parameters that are hard to identify have little effect on observed variables and therefore are probably not important.

(2) Manipulator links are not perfectly rigid.

When there is significant bending in the structure, the manipulator dynamics become much more complicated. In present-day industry there is a push towards lighter-weight manipulators to increase relative payload ability, speed of motion, and cost, but the price one pays is increased link flexibility. The underlying problem that has led to these developments in robot design is inadequate actuation with respect to power to weight ratio, especially when compared to human muscle. It is not clear this push would exist to the same extent if the actuation were better, and it does not seem that flexible-link dynamics is a particular source of worry in the biological system since bones do not bend very much.

In biological limbs the mass distribution can change due to muscle contraction; for example, the center of gravity of the thigh can shift by 10%. These changes are probably predictable, and while complicating the control problem do not pose the same level of complexity as flexible link dynamics. A potentially more significant problem is transmission flexibility, whether it be tendons, gear trains, or chains, which creates

[22] An on-line computation is one executed at the same time an associated process is running; an off-line computation takes place after the process has finished.

passive springiness at the joints. Again, transmission flexibility is less of a concern than link flexibility, since the rigid body dynamics still apply. Transmission flexibility may turn out to be an advantage rather than just a problem because, as mentioned earlier, it seems likely that some passive compliance will be needed for force control (Kazerooni, 1985).

(3) Actuator dynamics are not adequately taken into account or are too complex to model.

Due to nonlinearities in motors and amplifiers, control signals can bear a complex relationship to motor torque. Friction often provides an unpredictable element, arising from transmission elements or from intrinsic motor characteristics (Snyder, 1985). If actuator dynamics cannot be modelled usefully, they may dominate considerations of link dynamics. Nevertheless, actuator dynamics are still in some sense simpler because they are described by one variable compared to the n variables for link dynamics.

One way of compensating for an inability to model the actuation and transmission elements is to tune the output for specific movements through repetition. This approach is very reminiscent of the motor tape idea, in which the output is known only for one particular trajectory. According to this approach, general movements would be made coarsely or suboptimally with an imprecise system model and control, but for frequent movements the control system would modify its output for a new repetition based on errors from the previous repetition (Arimoto, Kawamura, and Miyazaki, 1985; Craig, 1984).

3.3.4 Feedback Control

However complete a dynamic model of a manipulator may be, it is not possible to predict exactly the actuator torques that will be required to execute a movement. There will always be some error in the model of the manipulator, and aspects of the model such as the actuator state may fluctuate. External disturbances that by their nature are not accounted for may also arise. In the human case, for example, putting on a coat perturbs arm movement. It is therefore considered essential that a feedback process exist to correct the inevitable errors in a trajectory. The inverse dynamics computation represents a feedforward process that attempts to predict the exact torques, and a feedback process works in conjunction with the feedforward process to correct the output. A feedback process is also necessary for force control, because the resultant motion is a consequence of the sensed contact force or kinematic errors.

In the most general form, the feedback law relating trajectory errors to corrective torques is cast in terms of task variables. After all, when controlling hand position the correction of errors is most sensibly done in hand coordinates rather than in joint coordinates, given the complex relation between joint errors and endpoint positions. It is interesting that studies of reflexes in humans indicate that corrections often occur, not in muscles, joints, or even limbs to which a perturbation is applied, but in remote sites that are appropriate for the motor tasks. Abbs and Gracco (1983) perturbed the lower lip during speech and observed upper lip compensation to maintain the speech goals. Similarly, Abbs, Gracco, and Cole (1984) perturbed a finger in a pinching task and found a compensation by the other finger. Lacquaniti and Soechting (1984) showed that reflex compensation at the elbow during perturbation of the whole arm is consistent with maintenance of joint torque rather than of any intrinsic muscle parameter. This separation of the response from the point of sensing is a necessary capability for achieving sophisticated control and argues against narrow reflexology.

A typical feedback law in hand coordinates is proportional-derivative (PD) control, where a position error is multiplied by a position gain and added to a velocity error multiplied by a velocity gain. The position gain is equivalent to a stiffness, and the velocity gain is equivalent to a damping. Other terms that may be added in this feedback law are desired acceleration (Luh, Walker, and Paul, 1980; Takase, 1977) and contact force (Hogan, 1984). The sum of all these terms yields a corrective hand acceleration that should be appropriate to reduce the errors. One must then convert hand acceleration to joint acceleration by solving the inverse kinematics, and then find the corrective torques by solving the inverse dynamics.

3.4 Biological Implications

This movement planning hierarchy represents a general motion control system, and illustrates the kinds of transformations that must occur explicitly or implicitly to realize a desired endpoint trajectory. An explicit realization would be a deliberate sequence of transformations as in the robotics model of a planning hierarchy, from a detailed point-by-point evolution of the endpoint positions, to the corresponding time sequence of joint angles, and then finally to the actuator torques required. An implicit realization would involve setting up some lower-level organization, perhaps muscle synergies for biological motor control or coupled joint activations, to evolve in such a manner as to approximate the movement goals.

We examine next how the movement planning hierarchy may be applied towards understanding human arm movement. First, what evidence is there for planning in

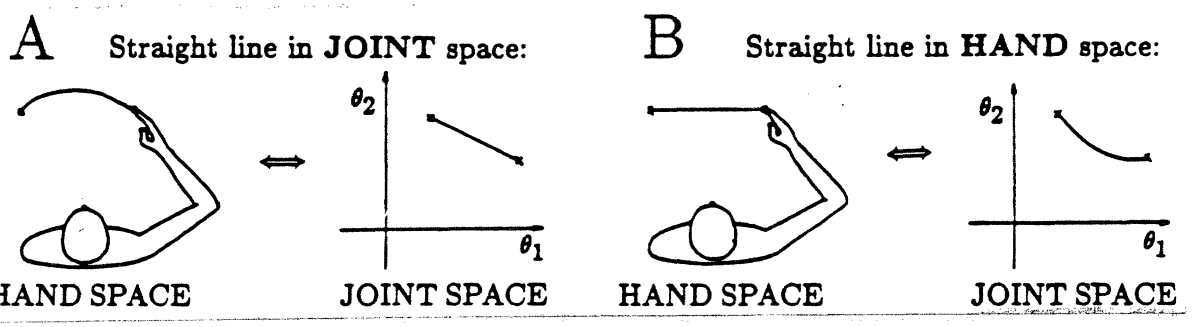


Figure 15: Different planning variables and their resultant trajectories for planar two-joint arm movement. (a) A straight line in joint coordinates generates a complex curved endpoint trajectory. (b) A straight line in Cartesian coordinates requires a relatively complex elbow and shoulder joint movement.

endpoint coordinates, as opposed to planning in joint coordinates or in actuator coordinates? Second, how can the motor control system be reconciled with movement dynamics?

3.4.1 Planning in Hand Coordinates

It would almost seem teleologically imperative that the motor control system have an ability to plan in terms of hand coordinates. The tasks of writing on a board, picking up and moving a cup, screwing in a lightbulb, and opening a door, given earlier as examples of external constraints, would seem to demand this ability. When planning in hand coordinates, the external constraints are most easily captured. The alternative of planning in more intrinsic coordinates presents the difficulty of how to predict the consequences of movement in the face of the complex transformation that take place between the various levels.

While planning in intrinsic coordinates intuitively possesses the easiest method to organize movement, this approach is viable only if simplifying strategies can be found that exhibit near-general, or at least adequate, behavior. Ordinarily one would expect that simple trajectories at one level should yield complex trajectories at another; for example, a straight line in hand coordinates yields a complex joint angle trajectory, while a straight line in joint coordinates yields a complex endpoint trajectory (Fig. 15). Said another way, there is a conservation of complexity in movement planning. With intrinsic planning coordinates, it must be explained how external constraints can be matched without requiring a controller more complicated than one operating at a higher level and doing the necessary transformations to lower levels.

Some experimental evidence in fact supports the concept of hand-coordinate plan-

ning, primarily straight-line trajectories in the act of making self-paced point-to-point reaching movements (Morasso, 1981). That Cartesian straight-line trajectories support hand-space planning is an argument based on Occam's razor: the simplest description of movement reflects how the movement is generated. This argument is similar to Bernstein's principle of equal simplicity (Whiting, 1984). If the spatial shape of a trajectory is invariant irrespective of the muscle scheme or the joint scheme, then the motor plan must be closely related to the topology of the trajectory and considerably removed from joints and muscles.

Hand-space planning has also been invoked to explain curved movements, such as those induced by requiring subjects to pass through a *via* point between start and goal (Abend, Bizzi, and Morasso, 1982). Flash (1982) found that modeling movement in terms of endpoint coordinates and requiring that these endpoint coordinates minimize jerk (the third derivative of position) captured the essential features of path shape and velocity profile.

If hand-space planning exists, then biological processes equivalent to inverse kinematics would have to exist as well, but there is no direct evidence of such processes. Soechting (1984) observed that in accurate pointing movements the wrist motion is only loosely coupled to the elbow and shoulder joint motion. Given the earlier discussion about spherical wrist joints and simplicity of the inverse kinematics solution, the experimental evidence is consistent with positioning being separated from orienting degrees of freedom in order to solve the inverse kinematics.

3.4.2 Planning in Joint Coordinates

Straight lines in joint angle space are known in robotics as *joint interpolation*, where all joints are executed in lockstep with the same time profile. The joint angles interpolate linearly from start to goal, and hence never reverse direction. Joint interpolation generates curved Cartesian paths for two-joint arm movement as shown in Figure 15A. Hence joint interpolation is an instance of planning in joint coordinates that does not generally allow one to realize simple endpoint trajectories.

If to circumvent the above limitations the definition of joint interpolation is generalized to allow one joint to start or finish before another, and a joint's time profile to expand or compress, approximately straight Cartesian trajectories can be generated in certain regions of the workspace (Fig. 16). This strategy is henceforth referred to as *staggered joint interpolation.* The affected workspace regions correspond to Cartesian straight-line motions where a joint is not required to reverse itself, since as mentioned above, joint reversal is not allowed in joint interpolation. When a joint must reverse

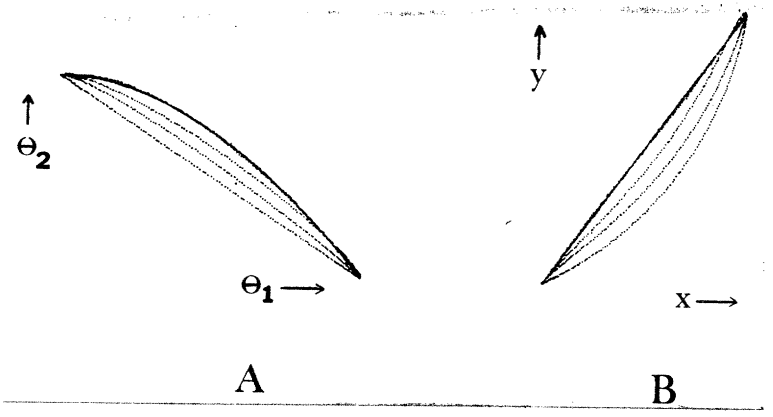


Figure 16: Joint angle plots of shoulder angle θ_1 versus elbow angle θ_2 (a) and corresponding endpoint trajectories (b) for perfect straight-line Cartesian trajectories (solid lines) versus staggered joint interpolation (dotted lines).

itself, a Cartesian straight-line path cannot be well approximated. The ability to generate approximately straight Cartesian paths by staggered joint interpolation cautions against automatically assuming that planning in hand coordinates is required to achieve straight Cartesian paths.

Recently evidence has appeared that in certain portions of the workspace human arm movements take on curved features explainable by joint interpolation. Corresponding to endpoint trajectories in Figure 17(a)-(d) between various targets in a vertical plane, the plots of joint angles in Figure 17(e)-(h) show that the curved trajectories (c) and (d) reflect straight lines in joint space. These movements correspond to workspace regions where joint reversal is required for Cartesian straight-line motion, and it is postulated that subjects in this experimental task refrain from joint reversal and adopt the simpler strategy of joint interpolation (Hollerbach, Moore, and Atkeson, 1985). Trajectory (a) is a special case of a Cartesian straight line passing through the shoulder, the only situation where joint interpolation generates a straight hand path. Although trajectory (b) is also approximately straight, it can be explained by staggered joint interpolation. In this instance the subject was able to find the best compromise to a Cartesian straight line by an appropriate choice of interpolation parameters.

A strategy demonstrably equivalent to joint interpolation has also been proposed by Soechting and Lacquaniti (1981), who found that in arm movements reaching towards the edge of the workspace the deceleratory phase consisted of a constant joint rate ratio between shoulder velocity and elbow velocity. Simulations based on this data are shown in Figure 18. In Figure 18a the plot of elbow joint velocity versus shoulder joint velocity for several trajectories shows an approach to a constant slope in the last half

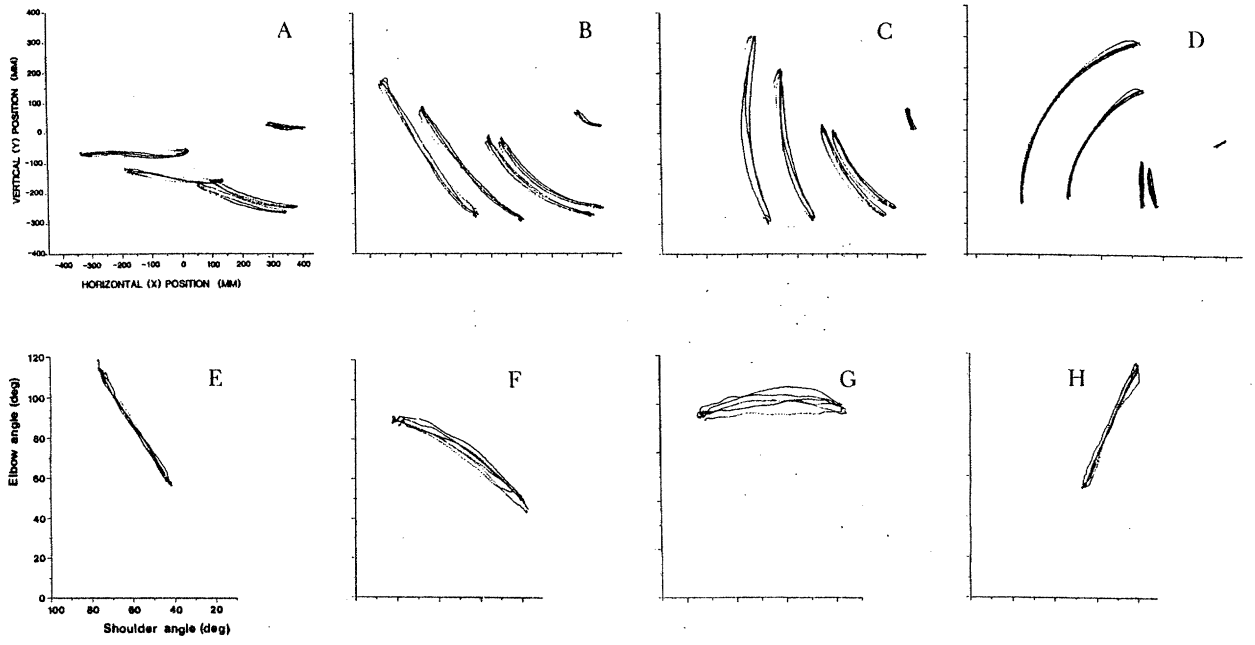


Figure 17: Trajectories of unrestrained arm movement between vertical plane targets measured with a Selspot system. The endpoint trajectories are shown in (a)-(d) as projected onto the vertical plane, and the corresponding joint angle plots of elbow versus shoulder angle are shown in (e)-(h).

of the movement. It has recently been shown, however, that any movement toward the workspace boundary approaches a constant joint rate ratio, regardless of the approach direction, location on the boundary, or coordination strategy (Hollerbach and Atkeson, 1985). In Figure 18b the movement plane is overlayed with contours of constant joint rate ratio. Movements of the endpoint in the lower right quadrant from the starting point towards various parts of the boundary traverse these contour lines to reach exactly the same joint rate ratio, which depends only on the link lengths and hence is a peculiar artifact of kinematics near the workspace boundary.

Thus the movements described in (Soechting and Lacquaniti, 1981) cannot by themselves be taken as evidence for joint interpolation, and a different set of experiments are required to make this argument that stay away from the workspace boundary. Although the initial part of the trajectories in (Soechting and Lacquaniti, 1981) did not show a constant joint rate ratio, it is nevertheless possible that the whole trajectory could be explained by staggered joint interpolation. The endpoint trajectories were relatively straight, but no joint was required to reverse itself. Once again, this analysis indicates that a superficial regularity at one level of description could have an explanatory underpinning at a different level.

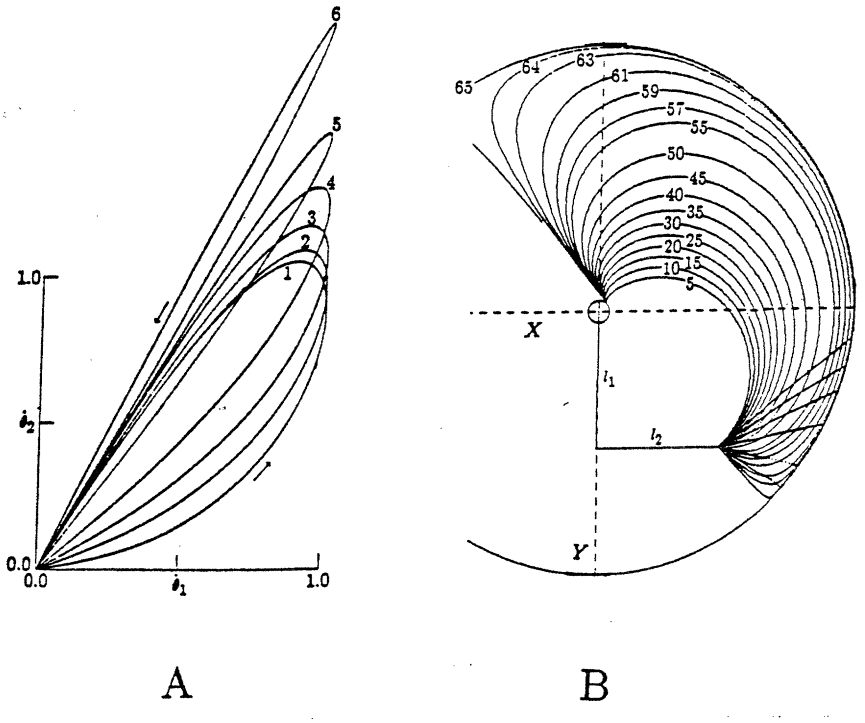


Figure 18: Simulations of vertical planar arm movements involving shoulder and elbow joints. In (b) the arm is shown in the starting position in the lower right quadrant of the movement plane. The center represents the shoulder point, and l_1 and l_2 are the upper arm and forearm lengths. The outer circle represents the workspace boundary, the points of maximal reach. The simulated movements begin from the starting position and approach different points on the boundary along straight-line paths. In (a) the ratio of elbow velocity to shoulder velocity are shown for each of these movements. Contour lines of constant joint rate ratio are imposed on the movement plane in (b).

3.4.3 Planning in Muscle Coordinates

Proposals have been made that utilize the viscoelastic properties of muscle to generate trajectories, and as such represent instances of planning in terms of muscle coordinates. Feldman (1974a, 1974b) and Crossman and Goodeve (1983) independently proposed the final position control hypothesis, also known as the equilibrium point control hypothesis. Muscles are assumed to act like springs to a first approximation, where for a given level of activation the force is proportional to length and a change in activation alters the spring stiffness or the zero setting. Around a joint the agonist and antagonist muscles act as opposing springs, and corresponding to their activations a total joint stiffness and equilibrium joint angle is defined that automatically generates restoring torques in response to a perturbation.

This essentially static model would be ideal to explain postural control, but the final position control hypothesis adapted this model to propose a basis for the generation of active movement. When the equilibrium point is shifted suddenly by changing the muscle activations to correspond to an equilibrium point at a desired final position, the mass attached to the springs will automatically move to and come to rest at the final equilibrium position. In this model, there is no explicit trajectory plan, which evolves dynamically through interaction of the moving mass with the potential field set up by the springs.

Some early experimental evidence supported the equilibrium point control hypothesis (Kelso and Holt, 1980; Polit and Bizzi, 1979), and theoretical models developed from it (Sakitt, 1980). More recent experiments employing perturbations, however, show the existence of intermediate equilibrium points (Bizzi, Chapple, and Hogan, 1982; Bizzi et al., 1984). Furthermore, theoretical simulation studies have convincingly shown that kinematic features of two-joint movements cannot be captured by this simple model (Delatizky, 1982). Hence the final position control hypothesis is now discounted.

3.4.4 Dynamics and Control

However one plans a trajectory, the correct joint torques and muscle activations must be arrived at to produce the movement. One of the most controversial issues in motor control is the extent to which the system knows about dynamics (Loeb, 1983). A main issue is the numerical computing ability of the nervous system, which it has been argued is inadequate to perform the many computer-like arithmetic operations required by even the efficient forms of inverse dynamics, acting at the servo rates that would be required during movement. Alternatives to analytic computation involve various forms of lookup tables for all or part of the dynamic equations (Albus, 1975a, 1975b; Raibert and Horn,

1978). These methods are similarly thought to require too much memory to represent with sufficient granularity the useful movement regions, even though these formulations were originally derived from theories of the cerebellum and motor learning (Albus, 1971, 1981; Raibert, 1978).

Instead of evaluation of exact nonlinear equations for dynamics, a frequently proposed alternative is a simplification or linearization of the dynamics, attempting to keep the most important terms and compensating the resultant errors with feedback. As mentioned earlier, it is unlikely that the biological system could implement this solution with proprioceptive feedback because of the transmission and processing delays. Two other alternatives are the utilization of viscoelastic muscle properties and the development of specially tuned movements, examined in more detail below.

Utilization of Muscle Properties

Because of muscle's viscoelastic properties, it is conceivable that an equivalent mechanical feedback could substitute for proprioceptive feedback. The equivalency results from the functional similarity of the actions of passive viscoelastic elements and active proportional-derivative (PD) control loops. Since viscoelasticity is a mechanical property of muscle, it acts instantaneously to resist perturbations and hence overcomes the basic speed limitations of active feedback.

The viscoelasticity of muscle is transferred to joints and ultimately to the endpoint due to the redundant musculature around joints and the presence of two-joint muscles. Whenever muscle contraction generates a joint torque or endpoint force, an apparent stiffness and viscosity is defined around the nominal state that will resist perturbations. As with active feedback, mechanical feedback will be most effective when the perturbations are not too large. If dynamics is treated as a perturbation, then ultimately the controllability of fast movements is limited (Johnson, 1982).

A scheme that proposes to treat dynamics as a perturbation through more explicit control of the effective endpoint viscoelasticity is the *reference trajectory hypothesis.* This hypothesis is derived from the final position control hypothesis in that it posits a sequence of equilibrium points from start to goal (Bizzi et al., 1984; Hogan, 1982). The multi-dimensional viscoelasticity of the endpoint can be set up around an equilibrium point to resist perturbations. The way the appropriate torques are generated is that the reference equilibrium point always moves in advance of the actual arm position, thereby creating a disequilibrium that propels the endpoint to follow the reference point. In effect, this strategy reduces movement to posture and dynamics to statics.

Research into the viability of this hypothesis is continuing; simulation results were

encouraging in capturing some detailed aspects of trajectories (Flash and Mussa-Ivaldi, 1984). A hypothetical equilibrium point trajectory was inferred from one measured trajectory and static stiffness fields, and was applied towards other workspace regions. The simulated trajectories captured the essential linearity of the corresponding experimental movements, even down to fine details of curvature. The extent to which the multidimensional viscoelasticity can be controlled is under study (Mussa-Ivaldi, Hogan, and Bizzi, 1984). One way this scheme could avoid problems with fast movements is if the viscoelastic properties scaled their intensities appropriately with movement speed to make the dynamics of the system time-invariant. It is not yet known if such is the case. It will also be necessary to demonstrate that the reference point moves in a simple manner and is invariant with speed; otherwise, it would just represent a different way of encoding dynamics.

Specially Tuned Movements

If the motor control system does not have a sufficiently accurate model of itself and if active or passive feedback processes cannot adequately compensate for dynamic motion, then the main alternative is specialized and individually tuned movements. Reminiscent of the motor tape idea, one is hesitant to propose this as an alternative because of the implied lack of flexibility. Nevertheless, thinking along these lines one would have to ask first if all movements are separately tuned or if there are elemental movements which serve as building blocks for more complex movements, second how these movements are actually tuned, and third whether decompositions exist that permit some flexibility in adapting to different conditions.

Currently little can be said one way or another about the existence of elemental movements. One possibility is that straight-line trajectories form a basic unit, which can be combined with some blending process to generate curved trajectories (Abend, Bizzi, and Morasso, 1982). Developmentally, it appears that babies adopt basic kinematic features of adult arm movement very early on (Fetters and Delatizky, 1984), so that perhaps these elementary movements are set up in early months and then slowly modified with growth.

With regard to tuning mechanisms, again not much is known about how this may come about, but perhaps the recent work in robotics mentioned earlier can serve as inspiration. The idea that movement regions could be represented coarsely or finely, depending upon the level of practice and skill, was explored in the context of robot dynamics by Albus (1975a, 1975b, 1981). This concept has been frequently mentioned as a possibility for motor control, e.g. (Loeb, 1983), but concrete proposals for how fine

vs. coarse tuning would take place are lacking.

Recently discovered time and load scaling properties of dynamics (Hollerbach and Flash, 1982; Atkeson and Hollerbach, 1985) could make the motor tape concept more attractive, because they allow flexibility with regard to changes of movement speed and hand-held load without requiring one to construct a completely new motor program. In Figure 19a the tangential velocity profiles of the wrist point for particular vertical-plane arm movements are normalized for time and distance to illustrate the underlying similarity in Figure 19b. It was found that these profiles were invariant for different trajectories, speed conditions, hand-held weights, and even subjects. The results were interpreted in terms of a massless phantom arm arm that carries the load and whose movement is superimposed on the physical arm (Figure 19c). By separately scaling the phantom and real arm for speed and load changes, through separation of the gravity torques from the inertial torques, simple linear combinations of these components were found to yield exact torque profiles for the different speeds and loads. In order for the scaling properties to simplify movement dynamics, the shape of the path and of the tangential velocity profile must remain invariant across speed and load changes, consistent with experimentally observed trajectories.

3.5 Conclusions

Biological motor control has been viewed from the perspective of a hierarchical planning and control structure derived from robotics. This perspective illuminates issues of kinematics, dynamics, and control that are an essential part of motor control but that are often overlooked in detailed physiological studies. The motion planning and control hierarchy represents a general-purpose structure that defines the transformations that must take place for the most advanced manifestations of movement control.

This general structure provides a framework for considering how the biological motor control system might derive its own solutions to the implied transformations. The basic question is how close to a general purpose structure is the biological motor controller? At the same time that limitations in control may restrict what can be accomplished, they may permit shortcuts in the transformations mentioned above. To answer the basic question requires much experimentation to determine exactly what are the bounds that circumscribe motor control. The lack of an adequate psychophysics of movement alluded to in the introduction creates a serious detriment towards progress on this issue. The search for regularities or invariances in movement production is an attempt to ferret out the motor control system's limitations. Many more experiments are required to test the extent to which these invariances hold or others might appear.

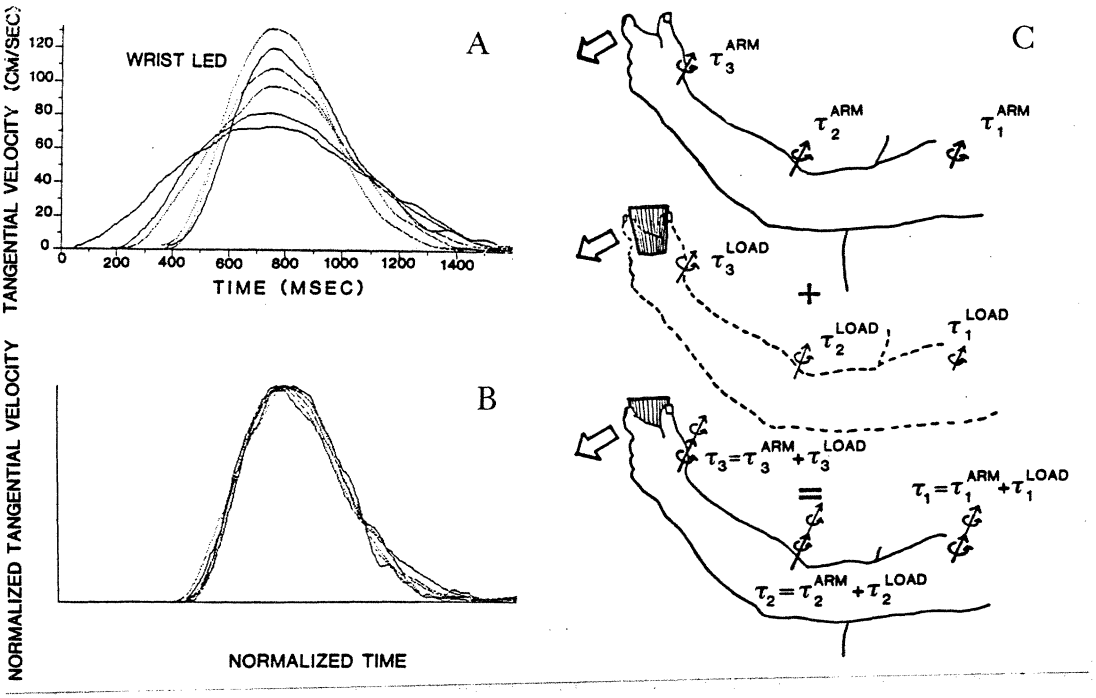


Figure 19: (a) The tangential velocity profiles of the wrist point for six vertical arm movements measured with the Selspot system. (b) The movements are normalized for time and distance to demonstrate the underlying invariance in profile shape. (c) A hypothetical phantom arm carrying the load and superimposed on the actual arm allows movement speed and load conditions to be simply changed if and only if the tangential velocity profile is invariant.

The computational approach to motor control is relatively new, and only a few investigators are applying the paradigm. So far there have been few concrete results. Indeed, progress in motor control research as a whole is moving rather slowly. A number of alternative biological strategies have been considered here, but it is premature to attempt to draw conclusions. The strategy of staggered joint interpolation may often allow a good approximation of Cartesian straight-line trajectories and would greatly simplify the inverse kinematics problem. The dynamic scaling properties of movement as implied by invariant tangential velocity profiles under different speed and load conditions show how a simple restriction on movement production could lead to simplification of the inverse dynamics computation. The viscoelastic properties of muscle could provide a feedback mechanism that avoids problems of transmission delay, and at the same time would unify position control and force control. Tuning mechanisms may exist to optimize certain movements that need to be repeatedly and accurately controlled, while leaving some more general purpose but coarser mechanism for less demanding movements.

What the computational approach offers is a fuller view of the scope of the motor control problem and ways in which it can be solved. It brings to bear the most recent advances in artificial intelligence, robotics, mechanical design, and control theory. Many of the general issues raised by the computational approach were already present in Bernstein's writings (Whiting, 1984), but technical advances have given better answers to old questions and raised new ones. Even since Saltzman's (1979) seminal paper on levels of sensorimotor representation, there have been significant advances in all aspects of control – trajectory planning, kinematics, dynamics, sensing, etc. – that have strong implications for motor control research.

The computational approach to motor control is intended to complement the research in motor psychophysics and physiology. Biologically specific constraints must be provided by experimentation, and hypothetical control strategies must be put to test. The hope is that the computational approach can contribute towards setting up more discerning experiments, interpreting data, and eventually discovering how the brain accomplishes its information processing tasks.

Acknowledgements: The authors wish to thank Chris Atkeson, Emilio Bizzi, Tomaso Poggio and Shimon Ullman, who provided valuable comments on this manuscript.

4 References

Abbs, J.H., and Gracco, V.L. Sensorimotor actions in the control of multi-movement speech gestures. *Trends Neurosci.* 6(9): 391–395, 1983.

Abbs, J.H., Gracco, V.L., and Cole, K.J. Control of multimovement coordination: sensorimotor mechanisms in speech motor programming. *J. Motor Behavior.* 16: 195–231, 1984.

Abend, W., Bizzi, E., and Morasso, P. Human arm trajectory formation. *Brain.* 105: 331–348, 1982.

Adelson, E. H., and Movshon, J. A. Phenomenal coherence of moving visual patterns. *Nature.* 300: 523–525, 1982.

Albus, J.S. A theory of cerebellar functions. *Math. Biosci.* 10: 25–61, 1971.

Albus, J.S. A new approach to manipulator control: the cerebellar model articulation controller (CMAC). *ASME J. Dynamic Systems, Meas., Control.* 97: 270–277, 1975.

Albus, J.S. Data storage in the cerebellar model articulation controller (CMAC). *ASME J. Dynamic Systems, Meas., Control.* 97: 228–233, 1975.

Albus, J.S. *Brains, Behavior, and Robotics.* Peterborough, N.H.: Byte Books, 1981.

An, C., Atkeson, C.G., and Hollerbach, J.M. Estimation of inertial parameters of rigid body links of manipulators. *Proc. IEEE Conf. Decision and Control.* Miami: December, 1985.

Anstis, S. M. The perception of apparent motion. *Philos. Trans. R. Soc. London Ser. B.* 290: 153–168, 1980.

Arbib, M.A., and Amari, S.–I. Sensori–motor transformations in the brain (with a critique of the tensor theory of cerebellum). *J. Theor. Biol.* 112: 123–155, 1985.

Arimoto, S., Kawamura, S., and Miyazaki, F. Can mechanical robots learn by themselves?. In: *Robotics Research: The Second International Symposium*, edited by H. Hanafusa and H. Inoue, Cambridge, MA: MIT Press, 1985, p. 127–134.

Atkeson, C.G., An, C., and Hollerbach, J.M. Rigid body load identification for manipulators. *Proc. IEEE Conf. Decision and Control.* Miami: December, 1985.

Atkeson, C.G., and Hollerbach, J.M. Kinematic features of unrestrained vertical arm movements. *J. Neurosci.* In press, 1985.

Ballard, D. H. Cortical connections and parallel processing: structure and function. *Behav. Brain Sci.* In press, 1985.

Barlow, H. B. Single units and sensation: a neuron doctrine for perceptual psychology?. *Percept.* 1: 371 394, 1972.

Barrow, H. G., and Tenenbaum, J. M. Recovering intrinsic scene characteristics from images. In: *Computer Vision Systems*, edited by A. R. Hanson and E. M. Riseman, New York: Academic Press, 1978.

Bishop, P. O., Coombs, J. S., and Henry, G. H. Responses to visual contours: spatio–temporal aspects of excitation in the receptive fields of simple striate neurons. *J. Physiol. London.* 219: 625–657, 1971.

Bizzi, E., Accornero, N., Chapple, W., and Hogan, N. Posture control and trajectory formation during arm movement. *J. Neurosci.* 4: 2738–2745, 1984.

Bizzi, E., Chapple, W., and Hogan, N. Mechanical properties of muscles: implications for motor control. *Trends Neurosci.* 5(11): 395–398, 1982.

Bizzi, E., Dev, P., Morasso, P., and Polit, A. Effect of load disturbances during centrally initiated movements. *J. Neurophysiol.* 41: 542–556, 1978.

Blakemore, C., and Campbell, F. W. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *J. Physiol. London.* 203: 237–160, 1969.

Bobick, A. A hybrid approach to structure from motion. *Proc. ACM Interdisciplinary Workshop on Motion: Representation and Perception.* Toronto: April, 1983, 91–109.

Borjesson, E., and von Hofsten, C. Visual perception of motion in depth: application of a vector model to three–dot motion patterns. *Percept. Psychophys.* 13: 169–179, 1973.

Boycott, B. B., and Wassle, H. The morphological types of ganglion cells of the domestic cat's retina. *J. Physiol. London.* 240: 397–419, 1974.

Brady, J.M., Hollerbach, J.M., Johnson, T.L., Lozano–Perez, T., and Mason, M.T., eds. *Robot Motion: Planning and Control.* Cambridge, MA: MIT Press, 1982.

Braunstein, M. L. Depth perception in rotation dot patterns: effects of numerosity and perspective. *J. Exp. Psychol.* 64: 415–420, 1962.

Braunstein, M. L. *Depth Perception Through Motion.* New York: Academic Press, 1976.

Breitmeyer, B. G., and Ganz, L. Temporal studies with flashed gratings: inferences about human transient and sustained channels. *Vision Res.* 17: 861–866, 1977.

Burt, P., and Sperling, G. Time, distance, and feature trade–offs in visual apparent motion.

Psychol. Review. 88: 171–195, 1981.

Campbell, F. W., and Robson, J. G. Applications of Fourier analysis to the visibility of gratings. *J. Physiol. London.* 197: 551–556, 1968.

Canny, J. F. Finding edges and lines in images. Technical Report 720 , MIT Artificial Intelligence Laboratory, Cambridge, MA, June, 1983.

Chomsky, N. *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press, 1965.

Cleland, B. G., Dubin, M. W., and Levick, W. R. Sustained and transient neurons in the cat's retina and LGN. *J. Physiol. London.* 217: 473–496, 1971.

Cleland, B. G., Levick, W. R., and Sanderson, K. J. Properties of sustained and transient ganglion cells in the cat retina. *J. Physiol. London.* 228: 649–680, 1973.

Clocksin, W. F. Perception of surface slant and edge labels from optical flow: a computational approach. *Percept.* 9: 253–269, 1980.

Cornsweet, T. N. *Visual Perception.* New York: Academic Press, 1970.

Cowan, J. D. Some remarks on channel bandwidths for visual contrast detection. *Neurosci. Res. Prog. Bull.* 15: 492–517, 1977.

Craig, J.J. Adaptive control of manipulators through repeated trials. *Proc. American Control Conf.* San Diego: June, 1984, 1566–1574.

Crossman, E.R.F.W., and Goodeve, P.J. Feedback control of hand–movement and Fitts' Law. Paper presented at the meeting of the Experimental Psychological Society, Oxford, July 1963, published in *Quarterly J. Exp. Psych.* 35A: 251–278, 1983.

Cynader, M., and Regan, D. Neurones in cat parastriate cortex sensitive to the direction of motion in three–dimensional space. *J. Physiol. London.* 274: 548–569, 1978.

Cynader, M., and Regan, D. Neurons in cat visual cortex tuned to the direction of motion in depth: effect of positional disparity. *Vision Res.* 22: 967–982, 1982.

Davis, L. A survey of edge detection techniques. *Comp. Graph. Image Proc.* 4: 248–270, 1975.

Delatizky, J. Final position control in simulated planar arm movements. *Neurosci. Abstr.* 8: 283, 1982.

DeMonasterio, F. M. Properties of concentrically organized X and Y ganglion cells of macaque retina. *J. Neurophysiol.* 41: 1394–1417, 1978.

DeMonasterio, F. M. Center and surround mechanisms of opponent color X and Y ganglion

cells of the retina of macaques. *J. Neurophysiol.* 41: 1418–1434, 1978.

DeMonasterio, F. M., and Gouras, P. Functional properties of ganglion cells of the rhesus monkey retina. *J. Physiol. London.* 251: 167–195, 1975.

Dreher, B., and Sanderson, K. J. Receptive field analysis: responses to moving visual contours by single lateral geniculate neurons in the cat. *J. Physiol. London.* 234: 95–118, 1973.

Enroth–Cugell, C., and Robson, J. G. The contrast sensitivity of retinal ganglion cells of the cat. *J. Physiol. London.* 187: 517–522, 1966.

Feldman, A.G. Change of muscle length due to shift of the equilibrium point of the muscle–load system. *Biofizika.* 19: 534–538, 1974.

Feldman, A.G. Control of muscle length. *Biofizika.* 19: 749–753, 1974.

Fennema, C. I., and Thompson, W. B. Velocity determination in scenes containing several moving objects. *Comp. Graph. Image Proc.* 9: 301–315, 1979.

Fetters, L, and Delatizky, J. Quantitative analysis of infant arm movements. *Neurosci. Abstr.* 10: 339, 1984.

Flash, T. Organizing principles underlying the formation of hand trajectories. Ph.D. Thesis, Harvard/MIT Division of Health Sciences and Technology, 1983.

Flash, T., and Mussa–Ivaldi, S.A. Inferring movement and muscle synergies from multi–joint arm posture. *Neurosci. Abstr.* 10: 635, 1984.

Gibson, J. J. *The Perception of the Visual World.* Boston: Houghton Mifflin, 1950.

Gibson, J. J. *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin, 1966.

Gibson, J. J. *The Ecological Approach to Visual Perception.* Boston: Houghton Mifflin, 1979.

Gibson, J. J. and Gibson, E. J. Continuous perceptive transformations and the perception of rigid motion. *J. Exp. Psychol.* 54: 129–138, 1957.

Goodwin, A. W., Henry, G. H., and Bishop, P. O. Direction selectivity of simple striate cells: properties and mechanism. *J. Neurophysiol.* 38: 1500–1523, 1975.

Graham, N. Visual detection of aperiodic spatial stimuli by probability summation amoung narrow band channels. *Vision Res.* 17: 637–652, 1977.

Graham, N., and Nachmias, J. Detection of grating patterns containing two spatial frequencies: a comparison of single–channel and multiple–channel models. *Vision Res.* 11: 251–259, 1971.

Green, B. F. Figure coherence in the kinetic depth effect. *J. Exp. Psychol.* 62: 272–282, 1961.

Grillner, S. Locomotion in vertebrates: central mechanisms and reflex interaction. *Physiol. Review.* 55: 247–304, 1975.

Grzywacz, N. and Hildreth, E. The incremental rigidity scheme for recovering structure from motion: position vs. velocity based formulations. AI Memo 845 , MIT Artificial Intelligence Laboratory, Cambridge, MA, August, 1985.

Hildreth, E. C. The detection of intensity changes by computer and biological vision systems. *Comp. Vision Graph. Image Proc.* 22: 1–27, 1983.

Hildreth, E. C. *The Measurement of Visual Motion.* Cambridge, MA: MIT Press, 1984.

Hochstein, S., and Shapley, R. M. Quantitative analysis of retinal ganglion cell classifications. *J. Physiol. London.* 262: 237–264, 1976.

Hochstein, S., and Spitzer, H. Zero–crossing detectors in primary visual cortex?. *Biol. Cybern.* 51: 195–199, 1984.

Hoffer, J.A., O'Donovan, M.J., Pratt, C.A., and Loeb, G.E. Discharge patterns of hindlimb motoneurons during normal cat locomotion. *Science.* 213: 466–468, 1981.

Hoffman, D. D., and Flinchbaugh, B. E. The interpretation of biological motion. *Biol. Cybern.* 42: 195–204, 1982.

Hoffman, K.–P., Stone, J., and Sherman, S. M. Relay of receptive field properties in dorsal lateral geniculate nucleus of the cat. *J. Neurophysiol.* 35: 518–531, 1972.

Hogan, N. Mechanical impedance control in assistive devices and manipulators. In: Brady et al., op. cit. 1982, p. 361–371.

Hogan, N. Impedance control of industrial robots. *Robotics and Computer–Integrated Manufacturing.* 1(1): 97–113, 1984.

Holland, H. C. *The Spiral After–Effect.* Oxford: Pergamon Press, 1965.

Hollerbach, J.M. A recursive Lagrangian formulation of manipulator dynamics and a comparative study of dynamics formulation complexity. *IEEE Trans. Systems, Man, Cybern.* SMC–10: 730–736, 1980.

Hollerbach, J.M. Computers, brains, and the control of movement. *Trends Neurosci.* 5(6): 189–192, 1982.

Hollerbach, J.M. Optimum kinematic design for a seven degree of freedom manipulator. In: *Robotics Research: The Second International Symposium*, edited by H. Hanafusa and H. Inoue, Cambridge, MA: MIT Press, 1985, p. 215–222.

Hollerbach, J.M., and Atkeson, C.G. Characterization of joint interpolated arm movements. *Exp. Brain Res. Suppl.* In press, 1985.

Hollerbach, J. M., and Flash, T. Dynamic interactions between limb segments during planar arm movement. *Biol. Cybern.* 44: 67–77, 1982.

Hollerbach, J.M., Moore, S.P., and Atkeson, C.G. Generalized joint interpolation explains straight and curved arm movements. *Neurosci. Abstr.* 11: in press, 1985.

Hollerbach, J. M., and Sahar, G. Wrist-partitioned inverse kinematic accelerations and manipulator dynamics. *Int. J. Robotics Research.* 2(4): 61–76, 1983.

Hollerbach, J.M., and Suh, K.C. Redundancy resolution of manipulators through torque minimization. *Proc. IEEE Int. Conf. Robotics and Automation.* St. Louis: March, 1985, 1016–1021.

Horn, B. K. P. Understanding image intensities. *Artif. Intell.* 8: 201–231, 1977.

Horn, B. K. P. *Robot Vision.* Cambridge, MA: MIT Press, 1985.

Horn, B. K. P., and Schunck, B. G. Determining optical flow. *Artif. Intell.* 17: 185–203, 1981.

Hubel, D. H., and Wiesel, T. N. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *J. Physiol. London.* 160: 106–154, 1962.

Hubel, D. H., and Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. London.* 195: 215–243, 1968.

Hulliger, M. The mammalian muscle spindle and its central control. *Rev. Physiol. Biochem. Pharmacol.* 101: 1–110, 1984.

Ikeda, H., and Wright, M. J. Receptive field organization of 'sustained' and 'transient' retinal ganglion cells which subserve different functional roles. *J. Physiol. London.* 227: 769–800, 1972.

Ikeda, H., and Wright, M. J. Spatial and temporal properties of "sustained" and "transient" neurons in area 17 of the cat's visual cortex. *Exp. Brain Res.* 22: 363–383, 1975.

Jacobsen, S.C., Wood, J.E., Knutti, D.F., and Biggers, K.B. The Utah/MIT dexterous hand: work in progress. *Int. J. Robotics Research.* 3(4): 21–50, 1984.

Jacobsen, S.C., Wood, J.E., Knutti, D.F., Biggers, K.B., and Iversen, E.K. The version I Utah/MIT dextrous hand. In: *Robotics Research: The Second International Symposium*, edited by H. Hanafusa and H. Inoue, Cambridge, MA: MIT Press, 1985, p. 301–308.

Jansson, G., and Johansson, G. Visual perception of bending motion. *Percept.* 2: 321–326, 1973.

Johansson, G. Perception of motion and changing form. *Scandinavian J. Psychol.* 5: 181–208, 1964.

Johansson, G. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14: 201–211, 1973.

Johansson, G. Visual motion perception. *Sci. Am.* 232: 76–88, 1975.

Johansson, G. Visual event perception. In: *Handbook of Sensory Physiology*, edited by R. Held, H. W. Leibowitz and H.–L. Teuber, Berlin: Springer–Verlag, 1978.

Johnson, T.L. Feedback control. In: Brady et al., op. cit. 1982, p. 127–146.

Kazerooni, H. A robust design method for impedance control of constrained dynamic systems., MIT, Mechanical Eng. February, 1985.

Kearney, R.E., and Hunter, I.W. System identification of human triceps surae stretch reflex dynamics. *Exp. Brain Res.* 51: 117–127, 1983.

Kelso, J.A.S., and Holt, K.G. Exploring a vibratory systems analysis of human movement production. *J. Neurophysiol.* 43: 1183–1196, 1980.

Koch, C., and Poggio, T. Biophysics of computation: neurons, synapses, and membranes. AI Memo 795 , MIT Artificial Intelligence Laboratory, Cambridge, MA, October, 1984.

Koch, C., and Poggio, T. The synaptic veto mechanism: does it underly direction and orientation selectivity in the visual cortex?. In: *Models of the Visual Cortex*, edited by D. Rose and V. Dobson, New York: John Wiley and Sons, 1985.

Kuffler, S. W. Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol.* 16: 37–68, 1953.

Kulikowski, J. J., and Tolhurst, D. J. Psychophysical evidence for sustained and transient channels in human vision. *J. Physiol. London.* 232: 149–163, 1973.

Lacquaniti, F., and Soechting, J.F. Behavior of the stretch reflex in a multi–jointed limb. *Brain Res.* 311: 161–166, 1984.

Land, E. H. Color vision and the natural image. *Proc. Nat. Acad. Sci.* 45: 115–129, 636–645, 1959.

Land, E. H. Experiments in color vision. *Sci. Am.* 200: 84–99, 1959.

Land, E. H., and McCann, J. J. Lightness and the retinex theory. *J. Opt. Soc. Am.* 61: 1–11, 1971.

Lappin, J. S., and Bell, H. H. The detection of coherence in moving random dot patterns. *Vision Res.* 16: 161–168, 1976.

Lathrop, R.L. Parallelism in manipulator dynamics. *Int. J. Robotics Research.* 4(2): 80–102, 1985.

Lawton, D. T. Processing translational motion sequences. *Comp. Vision Graph. Image Proc.* 22: 116–144, 1983.

Liegeois, A. Automatic supervisory control of the configuration and behavior of multibody mechanisms. *IEEE Trans. Systems, Man, Cybern.* SMC–7: 868–871, 1977.

Loeb, G.E. Finding common ground between robotics and physiology. *Trends Neurosci.* 6(6): 203–204, 1983.

Loeb, G.E. The control and responses of mammalian muscle spindles during normally executed motor tasks. In: *Exercise and Sport Sciences Reviews*, edited by R.L. Terjung, Lexington, MA: Collamore Press, 1984, vol. 12, p. 157–204.

Longuet–Higgins, H. C. A computer algorithm for reconstructing a scene from two projections. *Nature.* 293: 133–135, 1981.

Longuet–Higgins, H. C. and Prazdny, K. The interpretation of moving retinal images. *Proc. R. Soc. London Ser. B.* 208: 385–397, 1981.

Lozano–Perez, T. Task planning. In: Brady et al., op. cit. 1982, p. 473–498.

Luh, J.Y.S., Walker, M.W., and Paul, R.P.C. Resolved–acceleration control of mechanical manipulators. *IEEE Trans. Automatic Control.* AC–25(3): 468–474, 1980.

Marr, D. A theory of cerebellar cortex. *J. Physiology.* 202: 437–470, 1969.

Marr, D. Artificial Intelligence — a personal view. *Artif. Intell.* 9: 37–48, 1976.

Marr, D. *Vision.* San Francisco: W. H. Freeman Co., 1982.

Marr, D., and Hildreth, E. C. Theory of edge detection. *Proc. R. Soc. London Ser. B.* 207: 187–217, 1980.

Marr, D., and Poggio, T. From understanding computation to understanding neural circuitry. *Neurosci. Res. Prog. Bull.* 15: 470–488, 1977.

Marr, D., and Poggio, T. A computational theory of human stereo vision. *Proc. R. Soc. London Ser. B.* 204: 301–328, 1979.

Marr, D., and Ullman, S. Directional selectivity and its use in early visual processing. *Proc.*

R. Soc. London Ser. B. 211: 151–180, 1981.

Mason, M.T. Compliance. In: Brady et al., op. cit. 1982, p. 305–322.

Mayhew, J. E. W., and Frisby, J. P. Psychophysical and computational studies towards a theory of human stereopsis. *Artif. Intell.* 17: 349–385, 1981.

McMahon, T.A. *Muscles, Reflexes, and Locomotion.* Princeton University Press, 1984.

Minsky, M. *Semantic Information Processing.* Cambridge, MA: MIT Press, 1968.

Minsky, M., and Papert, S. *Perceptrons: an Introduction to Computational Geometry.* Cambridge, MA: MIT Press, 1969.

Morasso, P. Spatial control of arm movements. *Exp. Brain Res.* 42: 223–227, 1981.

Morgan, M. J., Moulden, B., Mather, G., and Watt, R. J. Intensity-response nonlinearities and the theory of edge localization. *Vision Res.* 24: 713–719, 1984.

Movshon, J. A., Adelson, E. H., Gizzi, M. S., and Newsome, W. T. The analysis of moving visual patterns. In: *Pattern Recognition Mechanisms*, edited by C. Chagas, R. Gattas and C. G. Gross, Rome: Vatican Press, 1985.

Mukerjee, A., and Ballard, D.H. Self-calibration in robot manipulators. *Proc. IEEE Conf. Robotics and Automation.* St. Louis: March, 1985, 1050–1057.

Mussa-Ivaldi, S.A., Hogan, N., and Bizzi, E. Invariant features of hand postural stiffness. *Neurosci. Abstr.* 10: 337, 1984.

Nagel, H.–H. Recent advances in image sequence analysis. *Proc. Premier Colloque Image.* Biarritz: 1984.

Olsen, H.B., and Bekey, G.A. Identification of parameters in models of robots with rotary joints. *Proc. IEEE Conf. Robotics and Automation.* St. Louis: March, 1985, 1045–1050.

Pantle, A. J., and Picciano, L. A multistable display: evidence for two separate motion systems in human vision. *Science.* 193: 500–502, 1976.

Peichl, L., and Wassle, H. Size, scatter and convergence of ganglion cell receptive field centers in the cat retina. *J. Physiol. London.* 291: 117–141, 1979.

Pettigrew, J. D., Nikara, T., and Bishop, P. O. Responses to moving slits by single units in cat striate cortex. *Exp. Brain Res.* 6: 373–390, 1968.

Pieper, D.L. The Kinematics of Manipulators under Computer Control. Ph.D. Thesis, Stanford University, Computer Science, 1968.

Poggio, G. F., and Poggio, T. The analysis of stereopsis. *Ann. Rev. Neurosci.* 7: 379–412, 1984.

Poggio, G. F., and Talbot, W. H. Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *J. Physiol. London.* 315: 469–492, 1981.

Poggio, T. Visual algorithms. In: *Physical and Biological Processing of Images*, edited by O. J. Braddick and A. C. Sleigh, Berlin: Springer-Verlag, 1983.

Poggio, T., and Koch, C. An analog model of computation for the ill-posed problems of early vision. AI Memo 783 , MIT Artificial Intelligence Laboratory, Cambridge, MA, May, 1984.

Poggio, T., Voorhees, H., and Yuille, A. A regularized solution to edge detection. AI Memo 833 , MIT Artificial Intelligence Laboratory, Cambridge, MA, May, 1985.

Polit, A., and Bizzi, E. Characteristics of motor programs underlying arm movements in monkeys. *J. Neurophysiol.* 42: 183–194, 1979.

Pratt, W. *Digital Image Processing.* New York: John Wiley and Sons, 1978.

Prazdny, K. Egomotion and relative depth map from optical flow. *Biol. Cybern.* 36: 87–102, 1980.

Prazdny, K. On the information in optical flows. *Comp. Vision Graph. Image Proc.* 22: 239–259, 1983.

Raibert, M.H. A model for sensorimotor control and learning. *Biol. Cybern.* 29: 29–36, 1978.

Raibert, M.H., ed. *Int. J. Robotics Research: Special Issue on Legged Locomotion.* 3(2): 2–146, 1984.

Raibert, M.H., and Craig, J.J. Hybrid position/force control of manipulators. In: Brady et al., op. cit. 1982, p. 419–428.

Raibert, M.H., and Horn, B.K.P. Configuration space control. *The Industrial Robot.* 69–73, June 1978.

Rajan, V.T. Planning of minimum-time trajectories for robot arms. *Proc. IEEE Int. Conf. Robotics and Automation.* St. Louis: March, 1985, 759–764.

Regan, D., Beverly, K., and Cynader, M. The visual perception of motion in depth. *Sci. Am.* 241: 136–151, 1979.

Richter, J., and Ullman, S. A model for the temporal properties of X-type and Y-type receptive fields in the primate retina. *Biol. Cybern.* 43: 127–145, 1982.

Richter, J., and Ullman, S. Simple cortical cells and the detection of zero crossings. *Submitted to Proc. R. Soc. London Ser. B.*, 1984.

Robinson, D.A. Models of the saccadic eye movement control system. *Kybernetik.* 14: 71–83, 1973.

Rodieck, R. W. Quantitative analysis of cat retinal ganglion cell responses to visual stimuli. *Vision Res.* 5: 583–601, 1965.

Rodieck, R. W., and Stone, J. Analysis of receptive fields of cat retinal ganglion cells. *J. Neurophysiol.* 28: 833–849, 1965.

Sahar, G., and Hollerbach, J.M. Planning of minimum-time trajectories for robot arms. *Proc. IEEE Int. Conf. Robotics and Automation.* St. Louis: March, 1985, 751–758.

Sakitt, B. A spring model and equivalent neural network for arm posture control. *Biol. Cybern.* 37: 227–234, 1980.

Salisbury, J.K., and Craig, J.J. Articulated hands: force control and kinematic issues. *Int. J. Robotics Research.* 1(1): 4–17, 1982.

Saltzman, E. Levels of sensorimotor representation. *J. Math. Psych.* 20: 91-163, 1979.

Sanes, J.N. Inaccurate limb positioning after unexpected perturbations during movement. *Neurosci. Abstr.* 9: 630, 1983.

Schiller, P. H. Central connections of the retinal ON and OFF pathways. *Nature.* 297: 580–583, 1982.

Schiller, P. H., Finlay, B. L., and Volman, S. F. Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *J. Neurophysiol.* 39: 1288–1319, 1976.

Schiller, P. H., and Malpeli, J. G. Properties and tectal projections of monkey retinal ganglion cells. *J. Neurophysiol.* 40: 428–445, 1977.

Shanmugam, K. S., Dickey, F. M., and Green, J. A. An optimal frequency domain filter for edge detection in digital pictures. *IEEE Trans. Patt. Anal. Mach. Intell.* PAMI–1: 37–49, 1979.

Shapley, R. M., and Victor, J. D. How the contrast gain control modifies the frequency responses of cat retinal ganglion cells. *J. Physiol. London.* 318: 161–179, 1981.

Sillito, A. M. The contribution of inhibitory mechanisms to the receptive field properties of neurons in the striate cortex of the cat. *J. Physiol. London.* 250: 305–329, 1975.

Sillito, A. M. Inhibitory processes underlying the directional specificity of simple, complex and

hypercomplex cells in the cat's visual cortex. *J. Physiol. London.* 271: 699–720, 1977.

Sillito, A. M., Kemp, J. A., Milson, J. A., and Berardi, N. A re-evaluation of the mechanism underlying simple cell orientation selectivity. *Brain Res.* 194: 517–520, 1980.

Slotine, J.–J. E. The robust control of robot manipulators. *Int. J. Robotics Research.* 4(2): 49–64, 1985.

Snyder, W.E. *Industrial Robots: Computer Interfacing and Control.* Englewood Cliffs, N.J.: Prentice–Hall, 1985.

Soechting, J.F. Effect of target size on spatial and temporal characteristics of a pointing movement in man. *Exp. Brain Res.* 54: 121–132, 1984.

Soechting, J.F., and Lacquaniti, F. Invariant characteristics of a pointing movement in man. *J. Neurosci.* 1: 710–720, 1981.

Spitzberg, R., and Richards, W. A. Broad band spatial filters in the human visual system. *Vision Res.* 15: 837–841, 1975.

Spitzer, H., and Hochstein, S. Simple– and complex–cell response dependencies on stimulation parameters. *J. Neurophysiol.* 53: 1244–1265, 1985.

Stein, R.B. What muscle variable(s) does the nervous system control in limb movements?. *Behav. Brain Sci.* 5: 535–577, 1982.

Takase, K. Task–oriented variable control of manipulator and its software servoing system. *Proc. IFAC Symp. Information–Control Problems in Manufacturing Technology.* Tokyo: October, 1977.

Thompson, W. B., and Barnard, S. T. Lower–level estimation and interpretation of visual motion. *IEEE Computer.* August: 20–28, 1981.

Tolhurst, D. J. Separate channels for the analysis of the shape and the movement of a moving visual stimulus. *J. Physiol. London.* 231: 385–402, 1973.

Tolhurst, D. J. Sustained and transient channels in human vision. *Vision Res.* 15: 1151–1155, 1975.

Torre, V., and Poggio, T. On edge detection. *IEEE Trans. Patt. Anal. Mach. Intell.* In press, 1985.

Tsai, R. Y., and Huang, T. S. Uniqueness and estimation of three–dimensional motion parameters of rigid objects with curved surfaces. Report R–921 , University of Illinois at Urbana–Champaign, Coordinated Science Laboratory, 1981.

Ullman, S. *The Interpretation of Visual Motion.* Cambridge, MA: MIT Press, 1979.

Ullman, S. Recent computational studies in the interpretation of structure and motion. In: *Human and Machine Vision*, edited by J. Beck, B. Hope and A. Rosenfeld, New York: Academic Press, 1983.

Ullman, S. Maximizing rigidity: the incremental recovery of 3-D structure from rigid and rubbery motion. *Percept.* 13: 255–274, 1984.

van Santen, J. P. H., and Sperling, G. Temporal covariance model of human motion perception. *J. Opt. Soc. Am.* 1: 451–473, 1984.

Victor, J. D., and Shapley, R. M. Receptive field mechanisms of cat X and Y ganglion cells. *J. Gen. Physiol.* 74: 275–298, 1979.

Wallach, H. On perceived identity: 1. The direction of motion of straight lines. In: *On Perception*, edited by H. Wallach, New York: Quadrangle, 1976.

Wallach, H., and O'Connell, D. N. The kinetic depth effect. *J. Exp. Psychol.* 45: 205–217, 1953.

Wallach, H., Weisz, A., and Adams, P. A. Circles and derived figures in rotation. *Am. J. Psychol.* 69: 48–59, 1956.

Watson, A. B., and Nachmias, J. Patterns of temporal interaction in the detection of gratings. *Vision Res.* 17: 893–902, 1977.

Watt, R. J., and Morgan, M. J. Mechanisms responsible for the assessment of visual location: theory and evidence. *Vision Res.* 23: 97–109, 1983.

Watt, R. J., and Morgan, M. J. The recognition and representation of edge blur: evidence for spatial primitives in human vision. *Vision Res.* 23: 1465–1477, 1983.

White, B. W., and Mueser, G. E. Accuracy in reconstructing the arrangement of elements generating kinetic depth displays. *J. Exp. Psychol.* 60: 1–11, 1960.

Whiting, H.T.A., ed. *Human Motor Actions: Bernstein Reassessed.* Amsterdam: North-Holland, 1984.

Wilson, H. R., and Bergen, J. R. A four mechanism model for threshold spatial vision. *Vision Res.* 19: 19–32, 1979.

Wood, J.E., Meek, S.G., and Jacobsen, S.C. Quantitation of human shoulder anatomy for prosthesis control. , University of Utah Center for Biomedical Design, October, 1984.

Yuille, A. L., and Poggio, T. Scaling theorems for zero-crossings. *Proc. IEEE Workshop on*

Computer Vision and Control. 1984, 3–7.

Yuille, A. L., and Poggio, T. Fingerprints theorems for zero-crossings. *Proc. Am. Assoc. Artif. Intell. Conf.* August, 1984.