ACTIVE PERCEPTION AND THE "RESONANCE METAPHOR"

Henrik Gedenryd

Lund University Cognitive Science
Kungshuset, Lundagård
S–223 50 Lund, Sweden
E-mail: Henrik.Gedenryd@fil.lu.se

This paper discusses active processes in perception and cognition, and introduces a "resonance metaphor," which emphasizes the active aspects of cognitive processes. In particular, active visual processes are discussed. Recently, the trend in computer vision has been a change from bottom-up approaches toward "active vision," much inspired by biology. Thereby the notion of active processes in biological (incl. human) perception has at long last been brought into computer vision. The resonance metaphor is demonstrated to be closely related to active perception, in contrast to a "conduit metaphor," which is closely related to information processing models. Finally, the relationship between active and non-active perception is related to "folk psychology" and eliminative materialism. In particular, it is argued that active perception is an example of scientific theories that eliminate concepts of folk psychology from scientific theories.

INTRODUCTION

During the recent five or so years there has been a shift in the directions of artificial (computer) vision. This shift was the result of a growing insight in the vision research community, an insight into the reasons for the general failure of artificial vision research until that time. That research was congruent with the general form of the artificial intelligence (AI) research program at that time, or what I will further refer to as *symbolic AI*. The new directions in vision have been labelled "active" vision. This paper first discusses the shortcomings of earlier artificial vision, and relates it to active perception. The "resonance" and "conduit" metaphors are then introduced, and their relations to active and "non-active" models of cognition are demonstrated . The resonance metaphor is then elaborated, and after that some examples of active influence in perceptual and cognitive processes are discussed. Finally, the philosophical and meta-theoretical implications of active perception and the resonance metaphor are discussed, in particular in relation to folk psychology and eliminative materialism. (Although in this paper I will mostly stick to vision, the principles will apply to perception in general.)

THE FAILURE OF TRADITIONAL COMPUTER VISION

What, then, are the problems associated with the traditional approach to vision? There are two major classes of failures on which the argumentation in this paper will be built:

1. *Identifying objects and features*. The first task in perception is to extract information from the sensory data; information that will guide actions. Old-school programs have only been able to identify objects under carefully controlled conditions: "…it is extraordinarily difficult to extract reliable line drawings in any sort of realistic cases of images… Real world images are not at all the clean things that our personal introspection tells us they are." (Brooks, 1991a, p. 576) For example, contours can practically never be extracted from an image—the fractional segments that can be found must be joined together by filling in the empty gaps between them.

2. *Maintaining and updating representations*. Typically, traditional approaches may take several minutes to create symbolic representations of the environment, and it is very hard to match objects between two successive versions of such representations. In a US. defense program the aim was to create a vehicle that would use vision to drive along a road. None of the projects using symbolic 3D models could manage to get anywhere near success, because there was too much error and deviation between the model's predictions and what actually happened (Ibid, pp. 578–579).

So why did the traditional vision systems fail on these two points? The first issue was the failure on extracting information from the visual stimuli. What these systems tried to do was to build a complete model of the environment *bottom-up*, with the model containing all entities that humans can recognize in a scene. These models were to be *objective*, that is, they should not be biased by what they would be used for in other modules. The fundamental principle of active systems is that they are built for a specific purpose;

Lund University Cognitive Studies – LUCS 22 1993. ISSN 1101–8453.

they are biased as to what they look for, so they *seek* certain things that will help them on their task. By doing this, they impose structure on the environment—the form of the constructed representation reflects the way it will be used. So in a way, these systems use expectations—either built into the system or established by the situational context—to *fill in* the things that are not easily inferred from the visual data.

The second issue was to maintain a working representation over time, achieving continuity over successive generations. Since earlier approaches were entirely bottom-up, they did not use any information from existing representations when creating new versions. Thereby the only guarantee for continuity was that the same things were found each time (except for minor differences caused by movements). This might have worked if only things were easy to identify, but this is not the case. Newer approaches on the other hand, use existing representations to seek information in the next generation of image data. By using those objects that have already been attributed to the environment (and therefore are in the representation) as expectations, continuity over time is obtained. In this way, "higher" cognitive processes come to *control* perception, and both temporal and spatial integration is accomplished. That is, we get continuity both in change over time and in spatial displacement. What active vision does is in a sense to *tightly couple* the external objects and features to their already existing representational counterparts, thereby actively "linking" them to each other in the perceptual processes.

THE FOUNDATIONAL SCIENTIFIC CONTEXT OF EARLIER COMPUTER VISION

As mentioned in the introduction, research on artificial vision started as a branch of the general symbolic AI research program (Newell and Simon, 1976). Excellent overviews of both topics can be found in (Brooks, 1991a, 1991b). In symbolic AI, the primary focus is on problem solving and similar highly abstract human activities. Perception and action are not really considered to be topics of cognition, but rather as in- and output units through which the reasoning mind is thought to connect with the physical world. As an "input function," perception is assumed to work by combining the sensory stimuli into successively more complex units in a hierarchical *bottom-up* procedure. The classical book on bottom-up computer vision is Marr (1982). The completed abstract units, generally known as *symbols*, will supposedly be fed into the cognitive (i.e., non-perceptual) modules. Metaphorically, symbols are mental units moving around in mental space. They have referents in the physical world, and perception serves to produce the adequate mental symbols by analyzing sensory data. This scheme follows the "conduit" metaphor (Reddy, 1979): information about physical objects in the environment "flows" into mental space through a perceptual tube (see figures 1–2).

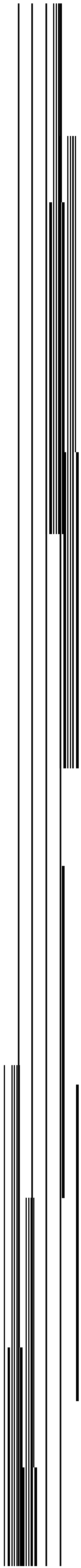

Figure 1. Traditionally, information-processing models of cognition consist of a sequence of modules. Perception is considered to be an input module, feeding information to higher levels in the form of symbols. Compare with figure 2. Adapted from Brooks (1991b, p. 1229).

This view is highly consistent with information theory (Shannon and Weaver, 1949), where a communicated message *carries* sequences of information objects, containing larger or smaller amounts of information (see figure 2). Computers came to play a very influential role in other sciences of the mind as well, e.g. psychology and linguistics, there known as "information processing models" of cognition (Pervin, 1989).

Other underlying assumptions of symbolic AI and information-processing models are that it is possible to describe the environment in a totally objective fashion, and that it really contains distinct entities—we do not ascribe distinct objects to the physical world, they are there, and they are objects independent of interpretations. This is why they can ”float” into the mind through a tube, without having to be created by mental processes. Perception merely converts them from physical objects to mental symbols. I will refer this as the notion of *objects in space*. Note that symbols are very similar to objects-in-space: they are distinct and have unambiguous, objective referents. Symbols therefore belong to the *objects-in-space metaphor* of mind; they are thought to reside in mental space.

THE SCIENTIFIC CONTEXT OF ACTIVE PERCEPTION

It is much harder to identify the more or less covert ideas of a contemporary research program than one that has been played out, or at least has lost its dominance as a scientific paradigm. Therefore, it is impossible to accurately describe the assumptions and implications behind active perception. The fundamental idea behind active vision is that perception works actively and selectively, instead of using simple bottom-up schemes. It is important to remember that these ideas about perception are not new. Helmholtz held such an opinion (see Grossberg, 1980), and Soviet psychologists had similar convictions quite early, see for example Luria (1973). Bruner carried these ideas quite far in the late 1950’s (Bruner, 1974). This only makes the earlier approaches to artificial vision seem even more mysterious. A general reorientation in AI towards biologically inspired work has certainly given birth to the reorientation in vision as well. Turning to biology has changed the focus from highly abstract and specifically human functions, such as problem solving and planning, to mimicking the functions of animals, surely because of hopes to reach higher goals in an incremental fashion. It is for example known that the frog’s vision works mainly by detecting dark spots in the visual field, allowing it to catch flies. Reproducing this artificially lies far from producing the general-purpose, viewer-independent models that were the goals of earlier attempts.

Active vision, and also active perception in general, contrasts bottom-up approaches by *creating* percepts by using sensory data. The consequence of this is that what we see is as much a product of the mind as it is a product of the physical world. It is no longer necessary for the environment to consist of distinct objects with clear-cut features, since interpreting perceptual mechanisms can *create* percepts from partial or ambiguous data. It is clear that such mechanisms produce *subjective* interpretations and not complete, general-purpose descriptions of the external world. It is similarly clear that the conduit metaphor is not applicable here; the information is not *in* the data, it is instead determined both by sensory data and by cognitive factors.

It is obviously inappropriate to connect active vision to the conduit and objects-in-space metaphors. Instead, as I will show, active perception is closely related to the *resonance* metaphor. The two major features of active perception are fill-in (or *reinforcement*) of weak or partial data, and tight coupling between external entities and their mental representations. These two are also the most salient characteristics of resonance: firstly, it amplifies and thereby sustains a vibration (e.g., a sound), and secondly, it causes two (or more) objects to vibrate in synchrony, so that changes in one object’s vibrations will alter the other object’s vibration as well. This latter influence serves to couple the two entities tightly to each other. (Compare these two points with the failures of earlier computer vision discussed above.)

In this paper I will focus on the first function of resonance, that is, reinforcement and fill-in, to demonstrate that perceptions contain contributions from both environmental and cerebral sources. I do this as a criticism of both symbolic AI and information-processing approaches to cognition, for their division of mental labor into distinct and separate input, processing, memory and output modules, etc. (see fig. 1).

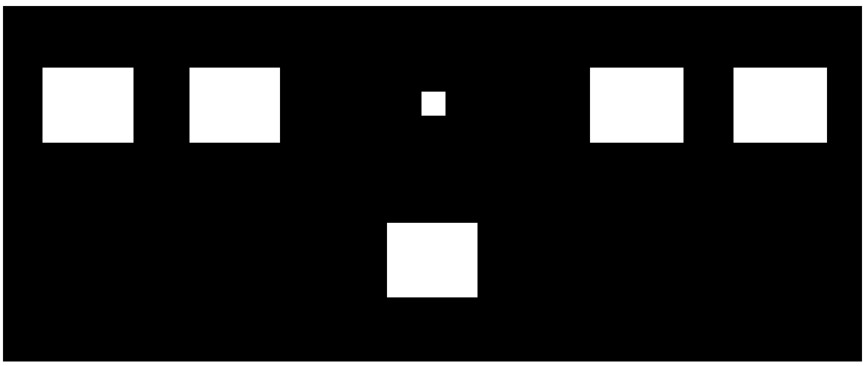

Figure 2. The connection between information theory and the conduit metaphor is quite obvious. Adapted from Shannon and Weaver (1949).

THE OBJECTS-IN-SPACE METAPHOR

A central theme in cognitive science is the role of metaphors in language. The study of metaphors is a linguistic revival of the studies of analogical thinking (Gentner and Stevens, 1983). An analogy uses a familiar concept to introduce a new one. In this way many aspects of the new concept need not be explicitly mentioned—they are inferred implicitly from the familiar, underlying concept. The metaphorical expressions that we use when talking about a certain concept are assumed to correspond to the underlying analogies by which we understand that concept. Different metaphors enable different ways of thinking about the same idea.

The words and expressions we use to talk about topics related to abstract psychological and cognitive topics indicate that we use an *objects-in-space metaphor*. For example, consciousness is usually spoken of in terms of a physical space in which objects reside (thoughts, feelings, etc.), at least in Western culture (Jaynes, 1976, as discussed in Roediger, 1979). Metaphors for phenomena related to memory follow the same pattern: memories can be objects in a house or in a cow's stomach (Roediger, 1979). Our notions of information and knowledge are further examples of physical space (Reddy, 1979). For instance, we regard learning as adding new material to memory, just like words are added to a dictionary, and books to a library. Even the very term "dictionary" is used in linguistics to refer to a place in memory where all the words that a person knows are stored. The area of computers and information processing uses the objects-in-space metaphor for information, which is evident in terms like "input," "output," "memory storage," "search" and "retrieval," etc.

Distinct entities and physical space seem to be involved in all these concepts, but generally on an indirect, abstract level. Hence I would like to call objects-in-space a meta-metaphor for cognitive concepts. I believe that it handicaps our understanding of certain aspects of perception, information and knowledge, by affording misleading inferences. It neglects the fact that the mind *actively* participates in the *construction* of experience. Instead, it regards experience as a stream of entities that flow from the environment into the mind through the senses (cf. fig. 1–2). Memory is conceived as a storage room from where we merely take out recollections. Entities are all fixed and ready as they enter awareness; whether they come from sensory data or memory. We *receive* perceptions—we do not *create* percepts from sensory stimuli. Similarly, we *retrieve* memories—we do not *reconstruct* memories out of remainders from previous mental activities.

THE RESONANCE METAPHOR

The resonance metaphor was originally applied to the area of perception (Gibson, 1966, Shepard, 1983), but it can easily be transferred to the areas of information and knowledge as well (Gärdenfors, 1992, ch. 7). This metaphor describes sensory data (visual and tactile stimuli, etc.) as *sound waves*. Bottom-up theories hold that a perceptual response is determined by sensory data only. Percepts are *in* the sensory data, and therefore float around in space, like sound waves do. The mind merely has to pick them up passively, just like a microphone picks up sound. The resonance metaphor contrasts this by stating that the mind works like a resonator, i. e. an object that has a "natural" or preferred state of vibration. This natural state depends on the object's own properties. Hence, according to the resonance metaphor percepts are not in the sensory data. Instead the mind actively *constructs* them; it begins to resonate.

The perceptual experience depends on factors in the mind itself, just like the resonance in a physical object depends on the object's own characteristics (see fig. 3).

Figure 3. Resonance depends on two factors.

The same pattern applies to memories as well. The way they are reconstructed depends on the specific circumstances in the mind at the time when they are remembered. A more general version of the resonance metaphor therefore does not only cover perception: no kind of experience depends only on the source that triggers it—instead, it is to a great extent affected by the mental status at that time.

PHYSICAL RESONANCE

The properties of a physical medium that affect its vibrations are called its *resonant characteristics*: a natural (or resonant) frequency and a tendency toward resonance. A stable, homogeneous object may resonate easily and powerfully while an unstable, heterogeneous one whose different parts vibrate in different ways may not resonate at all. Therefore, vibrations in an object with a low inclination towards resonance will depend only on external factors. A highly resonant object, in contrast, will itself to a great extent affect how it vibrates. External forces can hardly make it vibrate at a non-resonant frequency. On the other hand, even very weak external forces will induce a marked vibration if they only match its resonant characteristics. Therefore all externally induced vibrations in a resonant object will drift towards its natural vibration.

A tuning fork is specifically designed to take advantage of this phenomenon when tuning a musical instrument. If, for example, a piano string is out of tune, special interference effects, *drifts*, can be heard. When the tension in the vibrating piano string is changed, the drift will change character until it is in tune. Then the drifts disappear, and instead you hear a stronger tone than before. This happens because the interference effects that previously caused the two slightly different sounds to cancel each other out, now make the equal sounds reinforce each other. The important aspect for the metaphor is that an interaction between two factors determines the resonance: external forces and properties of the object itself. Two vital consequences follow: the resonant characteristics will transform any induced movement into a resonant vibration, and also amplify a weak vibration close to the resonant frequency (see table 1).

COGNITIVE INFLUENCE ON PERCEPTION

We compare perception to physical resonance to stress that some properties of the mind affect what we experience. What are these properties, then? Computer vision has shown that it is almost impossible to interpret the environment using only the visual image that reaches the eyes (or cameras). Put simply, we perceive things that the eyes do not see. A natural computer approach to vision attempts to find structures in the image and successively combine these into new ones of higher order. This is called a bottom-up approach: lines combine into eyes, a nose and a mouth, these together make up a face, which combines with hair, etc. into a head, and so forth.

The main obstacle in artificial vision is the low quality of the image data. Humans easily see contours where a computer only can find a number of small, quite unrelated line segments. No realistically simple algorithm can join such separate line segments into higher-order structures. We say that there is not enough information in the image; we mean that it is impossible to interpret it unless we have any clue to what it contains. You may have had similar experiences: at first being unable to read a sign (e. g., when an optician is examining your eyes). Then when you are told what it says, you can identify the letters. You do not simply *know* that they are there, it seems as if your eyes suddenly really *see* the letters that were obscured before. The clue activated the right mental structure, so that it could establish resonance with the visual data from the sign. In an interactive approach to vision, experience and context can give clues to the interpretation of an image. Where bottom-up vision only finds fragments without structure, other parts of the mind can provide clues for the visual processes as to what the data may contain. For example, the knowledge that you are watching football on television helps you in identifying a blurred, rapidly moving group of white dots as a football instead of something else that is inappropriate under these circumstances. Thus “higher-level” states (“I am watching football”) help to bring about and maintain “lower-level” states (“I see a moving football”).

There is an important distinction to make here: bottom-up strategies usually work well on computer input. “Real-world” images are different; they contain a great deal of superfluous data, “noise.” Computer input must be strictly delineated. It may be infinitely complicated, but it must follow a finite set of rules. This is normally not a problem, because the problems that programs are to solve are well-defined. In other words, the computer must do exactly what it is told, and only that. It has no other *raison d’etre*. Humans and other animals, on the other hand, have an important design feature that computers do not have; we survive, even without being told to do so. Perception is basically intended to keep us alive, so as long as we do stay alive, we may do whatever we wish with image data (Maturana and Varela, 1980). Knowing how to survive, we can attend to important parts of images, and ignore the rest. It is this knowledge that comprises the resonant characteristics of the mind.

Perception therefore does not pick up information from the environment through the senses by taking data and combining it into successively more complex structures in a top-down fashion. Instead it uses knowledge (in memory) to construct information that directs the processes involved in survival. Visual awareness comes out of combining dynamic image data on the retina with static mental patterns into experiences (see figure 3). Thereby memory contributes to perception.

RECONSTRUCTING MEMORY

It is relatively easy to accept that memory contributes to perception. It is more difficult to understand how recollections do not only depend on memory. We usually think of ourselves as lifting complete memories out of storage spaces in our minds. The distinction between recall and recognition is helpful here. Those of us who cannot draw a horse well, indeed say that we cannot *draw*. Instead I suggest that we cannot *recall* a horse. That is, we cannot retrieve a detailed image of a horse from the backs of our heads. If we just have a picture to copy, we do much better. Still,

Physical resonance	**Cognitive resonance**
• The vibration of an object is determined by two factors: external forces and the object’s own resonant characteristics.	• A percept is determined both by sensory data and by cognitive factors.
• The object’s influence over its own vibration will make it amplify and sustain a natural vibration in itself.	• The mental structures will cause a familiar perception to be amplified and sustained.
• Therefore any suitable external force, albeit weak and unstable, will produce distinct resonance.	• A familiar perception may be established from weak or unstable stimuli. (*fill-in*)
• The resonant characteristics will constantly pull any vibration toward the natural state of vibration. Therefore, any initial external force, how ever rough it may be, will drift toward resonance.	• The mental properties will strive to produce an as strong perception as possible. Hence, distorted or fragmentary stimuli will tend to be transformed in to a familiar perception. (*tight coupling*)

Table 1. Physical and cognitive resonance.

we recognize a horse just as well as a skilled artist does; we instantly see if some proportion is not right. Once we have drawn a horse from recall, we immediately reject it as ugly by recognition. Resonance explains this as a qualitative difference: recognition combines sensory data and memory into a distinct resonating experience, where recall lacks the details that awareness relies on the environment to supply. Resonance occurs, but there is no real external determinant.

An elegant experiment by Nickerson and Adams (1979, as discussed in Norman, 1988) has demonstrated that memories lack details: given several different drawings of American one cent coins, subjects were asked to pick the correct one (see figure 4). Recall was very low. Yet all of them certainly recognized real pennies and used them daily. Norman (1988) somewhat informally talks about "knowledge in the head" and "in the world," respectively. Normally we combine these two in daily behavior. An example of this would be when perception combines sensory data (from the world) with expectations or memories (in the head). For some reason we can use knowledge in the head without much help from the world (i. e., "remember"). I believe that this is a secondary ability; a side-effect, although a very useful one. There is more immediate ecological value in being able to apply experience to present situations, for example to "fill in" imperfect sensory stimuli, than being able to use experience by itself, as in recall.

HOW GENERALIZED MENTAL STRUCTURES STABILIZE PERCEPTUAL EXPERIENCE

I have now argued that neither percepts nor memories are independent of one another. We cannot experience either of them by themselves; experiences are always to some extent reconstructions which result from an interaction between sensory data and general experience. The adaptive nature of humans provides a natural explanation for this mixed quality of awareness. We can survive a very wide range of environmental conditions. To do this, we must be able to accommodate in ways that are not genetically accounted for. That is, we must be able to develop capabilities that none of our ancestors has ever had.

The principle that allows this kind of development is very simple: heritage only supplies a coarse nervous system, whose exact configuration is determined by how the different parts of it are used (Kandel and Hawkins, 1992, Shatz, 1992). (This, in turn, of course depends on the environment.) Often used portions develop at the expense of inactive parts. By letting the outcome of several situations influence the design of a subsystem, it will evolve by accumulating experience from past situations. Since the design is revised step by step, it will not correspond to any specific situation. Instead it will reflect a *generalized* situation. Aspects that reappear in several instances will stabilize in the design, and other aspects that do not follow any pattern will not translate into any lasting parallels in the resulting structure. Therefore, an incrementally revised structure will be general by nature; it will not contain any details.

The environment, on the other hand, contains constantly varying details. Successful adaptation depends on the ability to apply experience to new situations. Therefore, the varying details of the environment that affect the mind through sensory data must be matched with the general structures of experience. Necessarily a large number of detailed sensory patterns match with any general structure. The experience, deter-

Figure 4.

mined by both sensory data and mental structure, is similar although the details of the environment vary. The experience becomes an abstraction of the sensory patterns, and therefore indirectly an abstraction of the environment that caused them.

Such an experience is a "higher-order invariant" (Gibson, 1966, 1979; Shepard, 1983). Different aspects of the environment fluctuate in different ways. Those invariants that remain stable for longer periods come to seem more abstract than others. In this way the changes in the environment naturally translate into a hierarchy of invariants, where more abstract parts contain several others that are less abstract.

An *affordance* is a special kind of invariant. The theory of affordances is one of the most appealing parts of James Gibson's work (1979, ch. 8). Gibson's "realist" position does not allow him to assign affordances to individual experience (Ibid., pp. 129, 138–139), although he writes: "[an affordance] is equally a fact of the environment and a fact of behavior." (p. 129). This is exactly how a resonant experience works—it is determined both by environmental and cerebral elements. In resonance terms an invariant is a resonant experience that combines a stable mental structure with a varying aspect of sensory stimuli.

The mental structures provide another important stabilizing factor besides abstraction. It is reasonable to assume that environmental features that overlap will correspond to mental structures that also overlap. For example, some parts of the mental structures that correspond to a hand and an elbow will overlap the one that corresponds to the whole arm to which they belong. The mental structure representing the arm will supply parts of the momentum required to start resonance, and will therefore help to recognize the hand and elbow. It is this kind of influence that can account for the ability to read a sign only when knowing what it says. An important fact is that the overlapping structures affect each other *reciprocally*. That is, they influence each other in both directions, they *interact*.

In this context it is appropriate to discuss *categorical perception*, which occurs when "…within-category differences look much smaller than between-category differences even when they are of the same size physically." (Harnad, 1987, p. 535) For example in color perception, an equal difference in wavelength between two hues is perceived as larger when the two are categorized as different colors than when they are placed in the same color category. This phenomenon is a natural consequence of interactive perception. As previously stated, it is necessary for the mental structures to actively attempt to establish resonance with parts of the sensory data. The mental "resonators" must pull the sensory "vibrations" towards natural resonance in order to apply previous experience to the present situation. This is the need for categorization stated in terms of the resonance metaphor. The better the colors of perceived objects fit existing categories, the more they can be related to the individual's experience. It is therefore natural that resonators pull sensory color frequencies toward the centers of existing categories, away from the "gray" areas where they are of no use to the perceiver (Balkenius, 1992). I will not go into details of how mental resonance may be realized, but a detailed theory of mental resonance is Grossberg's ART (Adaptive Resonance Theory) (Carpenter and Grossberg, 1986, Grossberg, 1980).

PERCEPTUAL FILL-IN UNDER POOR CONDITIONS

Evidence of higher-level structures helping to determine what the lower-level information is, or even filling it in if it is missing, come from both perceptual experiments and computer simulations. Rumelhart and McClelland (1986, ch. 1) have simulated neural networks where higher-level units influence the selection of lower-level units, in what they call "an interactive activation model" (see figure 5). They used words and letters as the higher and lower levels, respectively. In that way, letters were selected if they fit into the word they were part of. Thus, context was used to judge which the individuals were most likely to be. Their model replicated the findings from earlier experiments on word-completion tasks by human readers (Ibid.).

Results similar to those from reading tasks have been found also in experiments on auditory perception. This has become well known as the *phonemic restoration effect*. This phenomenon amounts to that subjects hear sounds that have been cut out of recordings of speech, as if the removed sounds had been present. Not only did the subjects identify the higher-level structures (the words), they reported that they had *heard* the missing lower-level information (the phonemes) as well. (The reader is referred to Rumelhart and McClelland (1986, ch. 1) for details on such experiments in both visual and auditory perception.)

James Gibson claimed that the senses are only parts of a perceptual system, which also includes what the information-processing approach would have called the output from the mind, the motor system (Ibid., p. 42). He himself used the resonance metaphor to stress this view: “The peripheral retina registers such a locus, the brain resonates vaguely, and the eye is turned. Subjectively, we say that something ‘catches’ our attention. Then the orienting reactions begin and continue until the retino-neuro-muscular system achieves a state of ‘clearness’ and the brain resonates precisely.” (p. 260) This statement was a precursor of contemporary “active vision” approaches. However, his theory of direct perception holds that the senses can get information directly from the environment without inference, much in the same way as a stylus picks up music from a record (1979, ch. 14). Thus, Gibson holds that resonance is superfluous within the senses, but that it exists on the higher system level (i.e. between sensory and motor functions). Shepard has criticized Gibson on this point, saying that direct perception may work under favorable conditions, but that internal constraints “fill in the blanks” under inadequate conditions (Shepard, 1983, p. 422). He mentions three examples of such “reduced circumstances”: incomplete data, insufficient time and brain damage, of which I will only consider the first one. Shepard’s main argument against direct information pick-up in perception is that we are able to know of external entities that are unavailable to the senses, for example by being displaced in time or space, or by being hidden by darkness. Gibson probably proposed his theory because computer algorithms for pattern matching and extensive search seemed too slow to function as rapidly and smoothly as perception actually does. He was most certainly right on this point, but since then connectionist models fast enough to allow active perception have emerged. Nowadays we know through work on computer vision how often the sensory material is degraded, and this insight has made Gibson’s original theory obsolete (Brooks, 1991a).

MENTAL FILL-INS IN PERCEPTUAL ARTIFACTS

I will now discuss how a resonant model of perception, where memory plays a vital role, sheds new light on some perceptual phenomena. Theories that hold memory as separate from perception cannot readily account for these. According to them, percepts enter through the senses as finished, complete entities. Memory does not contribute actively, it is only used to identify the percepts. An interactive model, in contrast, gives memory an active role in the *creation* of perceptual experience. First, consider the perception

Figure 5. From “An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings” by J. L. McClelland and D. E. Rumelhart, 1981, *Psychological Review,* **88**, p. 380. Copyright 1981 by the American Psychological Association.

of ambiguous figures (fig. 6).

Figure 6.

According to the interactive model of perception, the unstable experience is caused by two alternate ways of creating information from the stimuli of an ambiguous picture. In figure 6 the lines are of course unambiguous, so interpreting them as lines is not a problem. The trouble arises when the lines are combined into steps (in 2 1/2 or 3D), because the lines can be interpreted as steps in two different, and conflicting, ways. Thus, both the conflicting interpretations of the individual steps may establish resonance, but only one complete staircase can be experienced. As the resonance process tries to stabilize itself, the two different interpretations may be equally strong, so the two complete interpretations will alternate. By concentrating on one of the two possible interpretations of the steps, the corresponding full structure will become stable.

The perception of non-existing shapes can be explained in a similar way. The cut-outs in the circles in figure 7 suggest the corners of a triangle, so they become “hints” that initiate a resonant process for a triangle. The triangle schema resounds by “expecting” sides of the triangle, and the sides are found in the triangles between the circles—perception has found a triangle. This phenomenon clearly demonstrates the active character of perception and the need to *assign* non-observable properties to the environment.

Figure 7.

A recent addition to the collection of perceptual curiosities is the red dot–green dot effect, also known as "color phi" (Dennett, 1991). When two dots are flashed after each other on a screen in slightly different positions, they are perceived as one moving dot. But if they have different colors, e. g. green and red, the perceived dot changes in color halfway, that is, "before" the second color is shown. The change in color is a side-effect of the effort to assign continuity to the two dots. The perceived dot changes its color to preserve the resonant link, and the mental inertia makes the transition smooth.

Finally, the interactive model of perception can account for hallucinations and mental imagery as resonant links that do not need external stimulation, see (Shepard, 1983). The interactive model does not build an image for some inner eye, but the experience arises since a mental structure resonates although the sensory data are missing. The virtual experience is just like "the real thing," because neither of these experiences depend on the details themselves, once resonance occurs. This fact can explain another perceptual curiosity. In this one, a visitor in an art museum enters a room where one wall is filled with paintings of Marilyn Monroe, all identical. The visitor experiences that he or she sees them all and that they all are alike. This happens although the eyes cannot see the paintings on the sides well enough to know for sure that they all are the same. The *experienced* image consists of the activated structures, and these do not depend on the details. Perception creates similar, equally good experiences where visual stimuli are not very good as where the eye records the finest details.

ACTIVE PERCEPTION AND INTUITIVE "FOLK PSYCHOLOGY"

A consequence of the resonance metaphor applied to vision is that what we as humans experience with our eyes is not "what is really out there," but rather something that vision creates to improve our fitness to the environment. If the resonance view on perception is the correct one, then vision works in a counter-intuitive way (and the other modalities do, too), because we experience our environment as real, and we also experience that our modalities are objective. That is, we do not experience that vision creates what we see. Instead we feel that we merely "pick up" information from the environment in an objective way. So why are we then unaware of that perception creates what we experience? Why do we think that the environment consists of well-defined "objects in space"?

At first, this seems to be a ridiculous question: it is almost an unquestionable axiom that the environment *is* a space with objects in it. The resonant model of perception claims that this is merely an interpretation, so how can it explain that we do not experience the interpretive processes? There is a straightforward explanation:

It is necessary for an adaptive organism to anticipate, that is, to be able to apply experience from previous events to new situations. In other words, it is quintessential to categorize and apply structure to every situation. This is equivalent to finding reoccurring properties in a changing system, and also abstract properties that remains constant although concrete, low-level properties fluctuate. The more familiar patterns you find in the environment, the more predictable it becomes—it becomes stable, since certain aspects reoccur and last over a longer time-span. So we become able to use more and more experience and expectations, and our fitness increases. *Therefore, it is a sign of success that we experience our environment as stable and distinct objects in space.* Evolution has brought us very far in making sense of our world. Our capacity to actively construct stable interpretations of our environment serves as the basis for our ability to act efficiently. So our experience of stable objects in space does not disconfirm the hypothesis of active perception. It is instead evidence for our ability to create our experiences in a highly adaptive way. Our cognitive functions are so successful in creating stable experiences that we do not realize that these experiences are fabrications, except for such rare cases as the perceptual curiosities discussed earlier.

The distinction between objective and active perception as described here fits into the discussion of *folk psychology* and *eliminative materialism*. Folk psychology is the label used for the naive or at least informal conceptions of psychology that people use to predict the behavior of individuals in their environment, for example. P. M. Churchland (1989) argues that these folk intuitions, which to a great extent have been incorporated into "scientific" psychological theory as well, will eventually be replaced as science progresses. He draws a parallel to "naive physics"; the intuitions of physics that people generally have, but which Newton showed to be false. Objects in space and objective perception clearly belong to the domain of folk psychology, since these are intuitive notions of how perception works. On the other hand, active perception as it has been described here is counterintuitive and therefore lies outside this domain. Its replacing of "naive" notions of perception would therefore be a good example of advances that eliminate from science folk intuitions about the workings of mind. More importantly, scientific areas that have adopted concepts from folk psychology, for example information processing models of cognition and symbolic AI, should realize the consequences of these advances. Specifically, these areas more or less implicitly assume that the mind operates on discrete symbols that designate discrete environmental entities. Active perception calls these assumptions into question, since the stability and discreteness of experienced entities are artifacts of the perceptual process.

Will Separate Mental Subsystems Disappear from Cognitive Theories?

If we are to speculate a bit more, you may remember that I initially criticized the sharp division between separate mental subsystems, such as units for perception, memory, action, planning, etc. (figure 1). I believe that our habits of dividing cognitive functions into highly disconnected parts may be artifacts of our own cognitive "strategies." As our minds try to understand each other (or themselves, or whatever), they apply the proven scheme of divide-and-conquer by chunking information into discrete entities, just like perception creates discrete "objects in space." Thus, to understand our mental systems, we collect their expressions into groups whose members appear to be similar. We thereby form categories such as *perception*, *memory*, etc. Further, we assume that these subclasses of behavior are performed by separate subsystems of cognition, although we have for a long time had reasons to believe differently. We have known for long that perception is not unmistakably different from cognition. For example, both go "beyond the information given" (Bruner, 1974) to make meaning out of data. By now we ought to recognize that the sharp distinction we make between these two systems is quite probably inappropriate. Therefore, as our views mind become more and more nuanced, the notion of separate cognitive subsystems ought to be dismissed from non-folk psychology.

Summary

Although it has been known for a long time that perception is active in character, and not at all an objective, bottom-up process, computer vision research has not until recently begun to acknowledge these facts. The reason for this has probably been the tight links between symbolic approaches to cognition and the notion of objects in space. Firstly, symbolic approaches implicitly assume that the world really consists of distinct entities. Secondly, this is also the intuitive "folk notion" of how perception works. As an alternative to objects-in-space, the resonance metaphor has been suggested. It neatly embraces the two major principles of active vision: perceptual fill-in and tight linking between stimulus data and mental representations. Further, some examples of visual fill-in have been presented. Finally, the relation between intuitive and "active" notions of perception, and the possibility of distinct cognitive subsystems being artifacts of active, constructive processes in cognition (in the same manner as active perception creates distinct entities) have been suggested.

REFERENCES

Balkenius, C., (1992), "Some properties of neural representations". In M. Bodén and L. Niklasson (ed.) *Selected Readings of the Swedish Conference on Connectionism 1992*, in press,

Brooks, R. A., (1991a), "Intelligence without reason". (ed.) *IJCAI'91—International Joint Conference on Artificial Intelligence*, pp. 569–595.

Brooks, R. A., (1991b), "New approaches to robotics", *Science*, **253**, pp. 1227–1232.

Bruner, J. S., (1974), *Beyond the Information Given*, George Allen & Unwin Ltd., London, England.

Carpenter, G. A. & S. Grossberg, (1986), "Neural dynamics of category learning and recognition: Attention, memory consolidation, and amnesia." In J. Davis, et al. (ed.) *Brain Structure, Learning and Memory*, Westview Press, Boulder, CO.

Churchland, P. M., (1989), *A Neurocomputational Perspective:the nature of mind and the structure of science*, MIT Press, Cambridge, MA.

Dennett, D. C., (1991), *Consciousness Explained*, Little, Brown & Co., Boston, MA.

Gärdenfors, P., (1992), *Blotta tanken (in Swedish)*, Nya Doxa, Nora, Sweden.

Gentner, D. & A. L. Stevens (ed.), (1983) *Mental Models*, Erlbaum Associates, Hillsdale, NJ.

Gibson, J. J., (1966), *The Senses Considered as Perceptual Systems*, Houghton Mifflin, Boston, MA.

Gibson, J. J., (1979), *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston, MA.

Grossberg, S., (1980), "How does a brain build a cognitive code?", *Psychological Review*, **87**, 51.

Harnad, S., (1987), "Category induction and representation". In S. Harnad (ed.) *Categorical Perception*, Cambridge UP, Cambridge, England.

Jaynes, J., (1976), *The Origin of Consciousness in the Breakdown of the Bicameral Brain.*, Houghton–Mifflin, Boston, MA.

Kandel, E. R. & R. D. Hawkins, (1992), "The biological basis of learning and individuality", *Scientific American*, **267**, pp. 53–60.

Luria, A. R., (1973), *The Working Brain—An Introduction to Neuropsychology*, Basic Books, London.

Marr, D., (1982), *Vision*, Freeman, New York, NY.

Maturana, H. R. & F. J. Varela, (1980), *Autopoiesis and Cognition*, D. Reidel Publishing Company, Dordrecht, Holland.

Newell, A. & H. A. Simon, (1976), "Computer science as empirical inquiry: symbols and search", *Communications of the ACM*, **19**, pp. 113–126.

Nickerson, R. S. & M. J. Adams, (1979), "Long-term memory for a common object", *Cognitive Psychology*, **11**, pp. 287–307.

Norman, D. A., (1988), *Design of Everyday Things*, Basic Books, New York, NY.

Pervin, L. A., (1989), *Personality: Theory and Research*, Wiley & Sons, New York, NY.

Reddy, M. J., (1979), "The conduit metaphor – a case of frame conflict in our language about language". In Ortony (ed.) *Metaphor and Thought*, Cambridge UP, Cambridge, England.

Roediger, H. L., III, (1979), "Memory metaphors in cognitive psychology", *Memory & Cognition*, **8**, pp. 231–246.

Rumelhart, D. E. & J. L. McClelland, (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, MIT Press, Cambridge, MA.

Shannon, C. E. & W. Weaver, (1949), *The Mathematical Theory of Communication*, Univ. of Illinois Press, Urbana, IL.

Shatz, C. J., (1992), "The developing brain", *Scientific American*, **267**, pp. 35–41.

Shepard, R. N., (1983), "Ecological constraints on internal representation: resonant kinematics of perceiving, imagining, thinking, and dreaming", *Psychological Review*, **91**, pp. 417–447.