Image Cover Sheet

1st CDN SYMPOSIUM ON REMOTE SENSING, 1972

PERCEPTUAL PRINCIPLES

RELATED TO REMOTE SENSING *

Dr. M.M. Taylor
Defence and Civil Institute of Environmental Medicine
P.O. Box 2000, DOWNSVIEW, Ontario.

SUMMARY

Remote sensing systems may be regarded as extensions of man's natural senses. As such, they should be governed by the same principles that govern natural perception. Perception is described as a means whereby information useful for action is separated from the enormous mass of useless information, and encoded in a way suited to rapid evaluation of potential behaviour. This functional viewpoint leads to the idea that the "attention" of a "central processor" must be devoted at any one time to a small region of the environment and that a behaving organism should be provided with a large number of feature detectors which continually monitor the environment for items that deserve the attention of the central processor and send "alarms" when such features are detected. Attention should continually shift except when called by these alarms. The vigilance decrement is perhaps due to an inappropriate requirement that attention be deployed on a single display for continuous periods of time, and might be averted by transforming the display in such a way that targets appear in a manner to which natural "alarm" detectors are suited. Similarly, the provision of hardware feature detectors should be a major part of remote sensing systems used in searches for known target types.

INTRODUCTION

Remote sensing systems are used in many ways and in many environments. Wherever man cannot use his natural senses, whether because the thing to be observed is in the wrong place, or because he is not sensitive to the kind of energy with which it must be observed, there exists a situation where remote sensing is valuable. Remote sensing systems are used in airborne geophysical surveys, to monitor city pollution, to search for lost aircraft, to monitor conditions inside a nuclear reactor, to detect invading aircraft. We map, we seek, we monitor, and we react to events indicated by the instruments. We may record data for possible later use, or we may want to respond as quickly as possible to critical events. Whatever the operation, except in the simplest, algorithmically controllable feedback systems, a human sits at the centre of the system. Sometimes many humans are involved in a system whose overall functioning is controlled by one person at the centre; sometimes there is only one man involved. But in most cases, the remote sensing systems gather and display their information so that a human can act. He may read a map, the product of a remote sensing system, and decide to despatch an oil drilling party, or he may look at a radar screen and decide to send an interceptor to investigate an unidentified aircraft.

In this paper, "remote sensing" refers to any method whereby a human gains information about the world other than through the unaided use of his natural senses. Remote sensing systems are extensions of the natural human senses, and their effective design should integrate with the design of the natural perceptual system. Indeed, even if a computer were to be the controlling centre for a remote sensing system, the same principles would apply, since they are based on the utilization of information rather than being inferred from results of experiments on natural perception.

Although current uses of remote sensing systems bear mainly on mapping functions, in which data is brought back for later selection by the human, this paper considers primarily situations in which the data should be used immediately if they are to be most effective. To this end, a functional viewpoint of perception is described, a viewpoint which suggests some possible ways to make remote sensing systems in conjunction with the human more effective. Although the viewpoint appears reasonable, its assertions are frequently derived from everyday experience rather than from formal study. For this reason, and because this is not intended as a paper for psychologists, the viewpoint is asserted rather than argued, and references are held to a minimum.

The Problem of Perception

Perception has evolved as a method of coordinating the behaviour of the individual. It is not just a method of making an image to match the world so that a homunculus can look at the image instead of looking at the world. Some theorists go so far as to identify perception with the total set of behaviour for which the individual could be prepared, but it is not necessary to take such an extreme position in order to recognise that the evolutionary success of perception must be due to the survival and reproduction of the perceiving individuals. In any behaving system, whether it is an organism or an army, effective behaviour depends on getting valuable information to the right place at the right time, as well as on having proper methods of dealing with the information and reacting to it. According to this viewpoint, the content of perception is the information currently deemed valuable and relevant to possible action. The individual or organisation perceives in order to act, and perception is closely tied to action possibilities.

The fundamental problem facing any perceiving system is that there is too much information available in the world, and too little of it is relevant to the needs of the moment. We are bombarded from all sides with electromagnetic radiation of all wavelengths; the air vibrates in all directions over a wide range of frequencies; surfaces react differently to chemical and mechanical probing; we touch and are touched by innumerable objects. We can increase without limit the number of properties we choose to test, and we can increase without limit the resolution with which we measure. The relative positions and motions of objects give way to the relations among their constituent molecules, and thence to atoms and the uncertain relationships among their elementary particles. Some information is in the structural patterns of things as they are at any given moment, some in the patterns of events as they occur.

In principle, one could analyse all this available information, relate the different structures, and extrapolate correctly the probable consequences of all possible actions. In practice, matters are quite otherwise. The central decision processor cannot possibly perform in one lifetime the analysis of even the static structures. As for the changes introduced by ongoing events, the information they introduce must often be acted on in seconds if the individual is to survive. The great problem is how to select the useful from the relatively useless information, and how to minimize the search for structural relationships whose potential number is very large for even a small amount of input data.

Natural perceiving systems have evolved a number of techniques to deal with the information overload problem. So obvious as to be often ignored is the complete and arbitrary elimination of most of the electromagnetic spectrum and most of the mechanical vibration spectrum. Similar arbitrary elimination of information is accomplished by limitation in the resolution of the sensors. It is technically possible to design an eye with any desired resolving power, and an ear with any desired sensitivity. But resolution and sensitivity beyond certain limits becomes less and less useful and more and more costly. The elimination of data sources unlikely to prove useful or likely to be too expensive is a good design principle for a perceiving system as well as for a National Science Policy. A remote sensing system is the result of a policy decision to use a data source eliminated by nature, and must be regarded as increasing the data reduction problem facing the human observer.

Information is structure. A reduction in the types of structure sought implies a reduction in information processing requirements. Most structure in the world is short range in terms of the sizes of elements composing the structure. Neighbouring points tend to have the same brightness, and when they do not, the same difference probably can also be found between a nearby pair of neighbours. Distant points rarely correlate well. Short time spans yield higher correlations than do longer ones. Accordingly, it is usually more profitable to look for short range structure, because more will be found. This is especially true when one considers that the number of possible twofold relationships in a two-dimensional space varies with the fourth power of the range over which they are sought, whereas the number of actual relationships over the longer ranges is probably less in total than the number of shorter range relationships. The neural cells needed to implement one simple pair-wise relationship among all pairs of retinal receptors would occupy a volume of some 300 cubic metres, at a conservative estimate. This figure alone shows how few of the possible structural relationships can actually be detected in any real system operating on natural data.

The cost-effectiveness of structural searches must decline very dramatically with increasing range. This fact is realized in the anatomy of natural perceiving systems, in that the elementary feature detectors typically have a high density of connection within a restricted neighbourhood and a very low connection density outside that neighbourhood. Even the so-called Fourier analyser elements may extend little more than a wavelength of the spatial frequency for which they are most sensitive. Larger-scale structure can readily be determined from the interrelations among the small-scale structures given by the primary feature detectors. In hearing, the probable relationships are among harmonically related features, and harmonic structure feature detectors should be common in the auditory system.

Wired feature detectors, at whatever level in the hierarchy, can respond only to the feature for which they are designed. The system designer has the option of incorporating wired detectors for all possible features or of eliminating unlikely features in favour of the ones that will most probably occur and be useful in practice. If the latter course is followed, as it must be in a device dealing with the complexity of nature, the system will be inflexible since the features discarded in the design may actually occur and be important on occasion. A system simplified by the elimination of possible fixed-feature detectors can be augmented by the addition of a variable, probably slow and non-algorithmic feature analyser which can look for new features of a kind not considered in the system design. This general analyser should then be capable in principle of discovering any kind of relationship that occurs in the environment. It should be used in most cases to determine complex or long range relationships, for relatively few of which would fixed feature detectors be economically feasible. The general analyser presumably forms part of the central processor which must decide on general courses of action on behalf of the entire individual, since it must take cognizance of relationships among effects from all data sources. It "thinks".

Information processing load reduction can thus be accomplished by ignoring possible data sources, and by ignoring possible relationships among data sources. The latter method of achieving reduction has a potentially severe difficulty in loss of adaptability, which can be mitigated by the incorporation of a general analyser capable of processing any relationships that might occur in the environment. Such a central processor pays for its capability with loss of speed. It also necessarily has to "pay attention" by considering at any moment only a small portion of the incoming data stream. This need to pay attention is the most significant aspect of the human central processor design from the viewpoint of the remote sensing system designer who wishes to provide the human with a useful instrument.

Interesting Things and Perceptual Modes

In order further to understand ways to accomplish information processing load reduction, we must consider how information is used in connection with behaviour. To this end, we can classify the things in the world that might be interesting in the sense that they might relate to current or subsequent behaviour.

Interesting things come in two major classes. These classes are not completely dichotomous, but shade into one another. Things in one class stay around long enough to be examined at leisure, while things in the other class are transient; they are here one moment and gone the next. What constitutes leisure depends rather on circumstances. A millionth of a second is a very long time in high-energy physics, but a month is short in the affairs of nations. Roughly speaking, if the thing being examined is there for a long time compared to the time taken to get all the essential data and to define the important relationships within the data, then it can be examined "at leisure". The processor can do many things before it has to pay attention to a stable thing. What constitutes leisure in the human time scale depends on circumstances, on what is being observed, and on what relations must be discovered in order to result in satisfactory behaviour. In ordinary perceptual experience, one would probably say that something which lasted only a few seconds was transitory, while something lasting a few minutes was stable.

Language usually classes things as "objects" or "states" on the one hand, and "events" on the other. Not all languages make the distinction in the same way. In English, there is a clear distinction between active events and stable states. The fact of change, however slow, is distinct from the state in any given time-frame. Objects, linguistically, have a permanent and individual existence, no matter how this permanency and individuality is belied by the real world.

In Hopi, according to Whorf (1940), the more prominent distinction is between conditions lasting longer or less long than a cloud. Fast things, like a lightning flash, a "setting of a stone", or a spoken command, are classed one way, whereas things slower to change or vanish, such as a house, a season, or a migration would be in the other class. The distinction between object and activity is less clear than in English, while the distinction between short and long-lasting phenomena is sharper. Both ways of speaking reflect valid views of reality. English focusses more on the relationship between actor and the thing acted upon, on the differences between the fact of change and the momentary state, whereas the Hopi focusses more on whether the condition must be observed now or whether observance can be put off till later. The Hopi view is more consistent with the emphasis placed here on the functions of perception which depend on whether or not the phenomenon will wait until the observer can attend to it.

The World Map

One of the major functions of perception is the construction of a world picture, a "map" of the stable states and objects in the world. The construction of this map is often taken to be the only function of perception, and its accuracy to be a measure of the adequacy of perception. Evolutionarily, this cannot be so. The adequacy of perception must be measured by the survival of the species, and the survival value of such a trait must be determined by its ability to induce correct action quickly enough to meet the circumstances. The production of a map is obviously one function of perception, but not the sole function. Neither is maximum map accuracy a reasonable measure of the value of the map. It must be an action-centred map, organized and arranged so that possible courses of action can be tested with its aid. Irrelevant detail is detrimental for speed. The map is more than a picture, but less than reality.

The map constructed by perception shares many qualities with a paper map. Both fall short of a complete representation of the region they depict. Both emphasize relationships that are useful for the purposes at hand. Just as a road map discriminates the surface qualities of the roads and may ignore the railways and the heights of land, while an agricultural map depicts fields of oats differently from fields of wheat, so the mental map must include coded representations of relationships that have been observed to exist in the world and are probably relevant to action. The labelling and coding by symbols and colours on paper maps are analogies of features which are often difficult to see in a photograph of the terrain. A skilled photo-interpreter must employ his skills before an aerial photograph can be turned into a map. Similarly, considerable coding and interpretation must be done before the raw sensory data stream can produce the relationships that go to make the mental map. The variability and complexity of these relationships suggest that they are largely discovered by the central processor. Indeed, since we have presumed that the central processor is responsible for the action decisions, it alone can know what features are currently relevant to action and should be incorporated in the map. This in turn implies the need for focussed attention in construction of the map. Map construction is relatively slow, partly because only a small part of the world can be scanned at one time, and partly so that the coding can be complete enough to permit fast map use when necessary for immediate action.

Neither the mental nor the paper map is used at the moment of its construction. Both serve as reduced and simplified sources of secondary information about portions of the world that may or may not be involved in later action. If action is required, however, the map makes it much easier to plan. Map data can be combined with current data, so that the terrain need not be resurveyed before each operation. The information is there and already correctly coded.

Attention and Alarm

Construction of the mental map requires focussed attention. Only a part of the map, and probably only a few features in that part, can be constructed at any one time. The world is too big and complex to be analysed all at once. Hence, the evaluative central processor must ignore most of the world most of the time. This does not mean to say that the whole individual ignores most of the world. The contrary is true. The wired feature detectors always operate on their portion of the data stream, regardless of where the attention of the central processor is deployed.

The focussing of attention implies a danger to the individual; events may happen and be missed if attention is not directed to them. This is the point of the distinction between things that can be examined at leisure and things that come and are gone. Attention can be deployed at will among the objects and stable states, but events must be examined while they happen or they will be missed. Of course, a certain amount of storage permits ongoing events to be held briefly in memory until they can be evaluated. Such storage would have to act like a shift register holding a certain duration of the entire data stream. Some storage of this kind is necessary, but it would be very costly to provide a long duration of such a wide-band shift-register. It would also not solve the problem of events that require fast responses. The simpler answer is to make a design decision that certain features probably signal situations demanding the attention of the central processor, and whenever these features are detected, an alarm calls the attention of the central processor to the region containing the alarming feature.

Consider vision. At first glance, vision seems to be the perfect example of a pattern processing system. Most remote sensing systems use visual displays and rely on the observer's vision to sort out the useful from the useless information. Yet little of the visual system itself is devoted to pattern processing. The whole system is a good example of the focussing and alarm principles on which attention is based.

When the eyes look straight ahead, one can see almost a full hemisphere. However, very little pattern is visible in the outer ranges of the hemisphere, and only the central 1^{o} is really useful for pattern vision. Acuity declines dramatically with distance from the centre. Twenty degrees from central vision the resolution has declined by a factor of ten, which means that the information transmission capacity for pattern has declined by a factor of 100. This decline is not matched by a parallel decline in the receptor density. Twenty degrees out, the receptor density is at least as high as it is in the centre, and this density provides the technical capability for the visual resolution to be as good in the one region as in the other. But the fibres of the visual nerve are nowhere near as plentiful 20^{o} out as at the centre, as if the potentially available resolution had been deliberately discarded, after having been built in. It looks at first sight like a design by committee, an unusual thing to find in a system tried and tempered by evolution.

What are all the extra receptors good for? In terms of alarm systems, the answer is apparent. The visual periphery is exquisitely sensitive to movement. Movement can readily be resolved between positions which cannot be resolved on the basis of static acuity. Even 20^{o} out from the fovea, a movement can be seen between points separated by only 3 minutes of arc, and possibly less. In the further periphery, very little can be seen at all unless something moves, and when something does move it is very obvious. It draws the attention. It "alarms".

Motion is a reasonable feature to have as an alarm in the visual system. Most of the visible world is fixed at any one moment, and movements usually signal either an event that must be accorded some immediate reaction. A good proportion of movements deserve attention, and those portions of the world that do not move can readily be scanned when time is available. Those movements that do not signal interesting events can be dismissed after a quick attention shift determines their uninteresting nature, and even if they continue they may not thereafter raise an effective alarm. Their uninteresting nature has been coded and can be determined for or by the central processor without an overt attention shift.

The ability to dismiss an alarm is a necessary function of an attention-getting procedure based on fallible analysis. Since the alarm system is likely to respond to a situation that does not really require the immediate attention of the central processor, and to continue responding to situations to which the central processor has already reacted, the central processor must be able selectively to switch off or to ignore an alarm.

Alarms in vision can be effective even when the detected feature is as simple as motion. In hearing, the situation is different. Hearing is not primarily a mapping sense. Most things we hear signal events, but this very fact suggests that most auditory events are not behaviourally significant and should not draw the attention of the central processor. In hearing, significant events must be distinguished from non-significant events by the peripheral preprocessors before an alarm is given. This means that there must be alarm feature detectors much more sophisticated than the simple motion detectors of the visual system. Furthermore, since simple waveform equivalence is inadequate for determining an event in the world, fairly complex and probably hierarchical sets of feature detectors must be involved. The fact that one's own name can serve as an alarm (as is shown by the ability to hear it in a noisy cafeteria) indicates that alarm features can be not only complex, but also learned. This is probably true also of alarms in modalities other than hearing. It suggests further that feature detectors themselves can be learned and do not have to be hard-wired genetically. For more complex features, most detectors are probably learned, probably in response to behavioural needs.

Implications for Remote Sensing Systems

The perceptual principles identified above carry several implications for the design of remote sensing systems whose data are to be used as they are gathered. They have little bearing on applications in which data are stored for later retrieval, as they are in mapping operations, for example.

The most immediate application is in the area of prolonged monitoring, or vigilance. The vigilance decrement is a term used to describe the inability of an observer to maintain his efficiency in detecting critical but infrequent events. The vigilance decrement seems to be a product of the technological age, and is usually associated with the need to pay continuous attention to a data source that does not provide much information and requires little reaction. According to the viewpoint on perception presented here, this is a most unnatural situation. Unless some event that needs immediate reaction or which requires a map update is going on, attention should not be devoted to a single data source. The primary task of the central processor is control of behaviour, which specifically requires continual restructuring of the world map. This may be accomplished by deploying attention among the various possible data sources, or by thinking about structures inherent in data already obtained. Mapping is prevented by paying attention to an uninformative data source. It seems that the conflict between the task requirements and the requirements of map updating is usually resolved by reducing the attention paid to the task. It is only fair to point out that this is a speculative hypothesis about the vigilance decrement, for which there is as yet no satisfactory published theory.

If this hypothesis concerning the cause of the vigilance decrement has merit, then remote sensing system designers, including the designers of radar and sonar systems, should beware of giving an operator an information-free and behaviour-free task. The operator should be allowed ample time to attend to matters other than the data display. Ideally, the possible presence of an interesting target should be signalled to the operator's alarm system. Motion perceived in the visual periphery would probably be adequate. A possible scheme to accomplish this in an early-warning radar involves time-compression. During each second of real time, the display would show a speeded history of the last period, say one minute, of the incoming data. Intruding aircraft would show as streaking lines, and noise clutter would show characteristically different motions. The streaks in motion associated with possible targets should be quickly seen, even by an operator occupied with a different task, provided that the screen was in his visual field.

Displays incorporating an alarm function, such as the time-compressed radar display, should have some facility for removing the alarm capability from an identified target. In the case of the radar display, for example, it is likely that the display would be digitally generated, and that a computer would be at the heart of the display. The operator should be able to identify to the computer the fact that he had detected the target, and the computer should react by computing the probable future track of that target. So long as it maintained a track near to that predicted from the history prior to the operator's response, the display should be kept dim, but if it deviated from expectation, the track should brighten until the operator again reacted to it. Permitting the operator to delete the alarm function from a detected target should enhance his ability to pick up new targets, while at the same time allowing him to keep track of the old. As was pointed out above, a necessary concomitant of an alarm system is the ability to turn off the alarm, and if the hardware does not permit this, then the operator's central processor will probably do it internally, so that new targets on the same display will be ignored.

A second difficulty with displays which rarely change their content of interesting information is that for behavioural purposes they become unchanging, with the consequence that it is more efficient for the central processor to work with the already coded map image of the display than with the incoming sensory data that needs coding. Reversion to memory mode may happen without the awareness of the individual, as witness the driver who goes through a new stop sign on a familiar route without ever seeing it. A driver to whom the route was new would always see that same stop sign. This phenomenon is well known to the police, who usually station an officer near every new stop sign. Inappropriate use of this memory image mode of perception can be reduced by changing the information content of the display periodically, or by removing it from the observer's field of view for long enough periods that he cannot rely on his memory to reconstitute it. This technique is probably rarely effective in itself, but the possibility should be considered in a design when unchanging displays can happen.

A third consequence of the analysis of the perceptual system is the suggestion that there should be a major role in remote sensing systems for hardware feature detectors of any degree of sophistication, so long as they have a different probability of responding to targets and to non-targets. On the other hand, the analysis also suggests that attempts to provide fully automatic pattern recognition devices to discriminate any but the simplest targets in the simplest contexts are doomed to failure. The main function of hardware preprocessors should be to provide a fast, wideband and wide-field scan of the incoming data so that the preprocessor can identify portions of the data stream worthy of the human observer's attention. In a search task, for example, any reduction in data load is welcome. Preprocessors such as linear contour detectors might be useful in enhancing the likelihood that the observer would detect cultural artifacts. Man-made objects are more likely to be bounded by straight lines than are natural variations in intensity or colour. If a visual display were made to contain only the outputs of line detectors which covered the whole field, but which could be switched by the operator at will to a full-picture mode, the operator would be more likely to detect man-made objects in the wild than he would from a continuous full-picture display. Thought should also be given the possibility of using the good learning capabilities of the human auditory feature detectors.

The major principle involved in the use of feature detectors is that the observer should not have to attend to the incoming data unless there is a reasonable probability that it contains something interesting. But when such an occasion arises, the full data stream should be available to him. The feature detectors signal where to look; they do not presume to substitute for the pattern recognition skills of the man.

Feature detectors, or simple pattern recognition devices, can be seen to be very important in the design of remote sensing systems for real-time action. They are probably also useful in the slower scanning of imagery recorded during planned survey missions, provided that there is a definable distinction between interesting and uninteresting things in the imagery.

CONCLUSION

The principles of perception, seen from the viewpoint of information processing load, provide a useful guide to the design of remote sensing systems intended for human operators. It is not necessary or desirable to flood the operator with as detailed a picture of the world as the sensors permit. But when the world contains something the operator should see, the system should warn him, and permit him to see it in as much detail as possible. The remote system designer should neither overload the operator with information to be processed nor force him to pay attention to a single data source for long continuous periods of time.

REFERENCES

(1) B.L. Whorf. Science and Linguistics. Technol. Rev. 42, 1940. 229-231, 247-248. Reprinted in Whorf B. L. Language, Thought and Reality, 1956, MIT Press.

* DCIEM Research Paper No. 828

NO. OF COPIES NOMBRE DE COPIES	COPY NO. COPIE N°	INFORMATION SCIENTIST'S INITIALS INITIALES DE L'AGENT D'INFORMATION SCIENTIFIQUE
1	1	
AQUISITION ROUTE FOURNI PAR ▶	DCIEM	
DATE ▶	29 Feb 96	
DSIS ACCESSION NO. NUMÉRO DSIS ▶		

DND 1158 (6-87)

National Defence Défense nationale

SN 228777