*Perceptron Applications in Photo Interpretation**

ALBERT E. MURRAY, *Principal Research Engineer, Cognitive Systems Grp. Physics Div., Cornell Aeronautical Laboratory, Inc., Buffalo* 21, *N. Y.*

ABSTRACT: *Perceptron is a general name for a large family of perceiving and/or recognizing automata. Many of these, when given suitable sensory equipment. resembling an eye, and consisting of a lens and a retinal mosaic of light-sensitiz'e elements, may be "taught" to reliably recognize or classify simple visual patterns. A laboratory model of a small and relatively simple perceptron has been used in experiments to estimate its potential in photo interpretation. In this program, it demonstrated an ability to learn, from a few examples, to visually identify simple, compact, well-defined objects, either alone or in company with other objects. The test objects successfully used ranged from alphabet letters through simple geomet'ric figures to simplified images from aerial photographs. Superimposed images of other well-defined objects or of relatively disorganized backgrounds do not seriously degrade the performance except in very special or severe cases. A plan is outlined fo·r further work toward development of an automatic photo-classification machine.*

PERCEPTRON (ref. 1, 2) is a general class name for a potentially large and diversified family of machines capable of perceiving and recognizing patterns. According to a recent definition of the term by its originator^a it is a convenient name to encompass all of the neural net type of "intelligent" machines. However, in this paper, the concern is primarily with the simpler mechanisms, particularly those which have been described in the literature under this name. The intent of this paper to to discuss preliminary experiments relating to photo interpretation, rerecently completed^b on a laboratory model perceptron and to consider some of the implications and probable extensions of the work.

The machine on which the experiments were performed, called the MARK I Perceptron, can be taught to recognize simple graphic forms. It tends strongly to generalize

(Tentative Title: Neurodynamics-Perceptrons tion. and the Theory of Neural Nets). The experiments which will be discussed

design and conduction of these experiments. Work supported by Office of Naval Research, Contract were deemed to be typical of, or fundamental No. Nonr-3161(00), Project NR 387-026. to, those required in the interpretation of

from a sample of training forms to give correct, identifying responses to any new forms which are similar to those in which it had experience during training. By suitable training on a few examples, the machine can also be taught to separate or distinguish forms which are generally similar but which possess certain kinds of distinguishing characteristic(s) .

The MARK I, which is representative of all elementary perceptrons, was not designed for specific tasks such as photo interpretation, but rather as a minimum model for some portion of the (visual) perception mechanism of a brain. It represents the simplest organization, with the smallest number of detailed specifications, of neuronlike elements, that might be expected to exhibit brain-like properties. Future photo interpretation research in our laboratory will employ machines which are related to perceptrons but whose development is aimed spe- ^a F. Rosenblatt, in C. A. L. Report VG-1196-G-8 cifically at applications in photo interpreta-
Centative Title: Neurodynamics—Perceptrons tion.

^b Mr. John Hay participated substantially in the herein were concerned with operations which

* Presented at the Society's 27th Annual Meeting, The Shoreham Hotel, Washington, D.C., March 19-22, 1961.

ALBERT E. MURRAY

aerial photographs. **It** is hoped that this continuing program may lead to the development of machines capable of performing a limited number of tasks which are selected for the purpose of reducing the burden which large quantities of survey data would place on trained personnel. A simple but worthy automation goal would be the ability to scan large quantities of mostly uninteresting photographic material in search of only one kind (or a small number of kinds) of object or feature, or one class of topography or terrain. The ability to recognize simple target" forms in relatively featureless environments is not a trivial task for a machine, but it is a capability which should be inherent in the class of perceptual mechanisms to which the MARK I Perceptron belongs. Consequently, the experiments reported herein begin with this problem, at a very formal and idealized level. They advance, through a series of tests of ascending difficulty, toward the recognition of practically interesting target forms in real and complex surroundings.

^c The term "target" as used in this paper, is intended to replace the phrase "one (kind of) object, feature, topography or terrain" and is to indicate the object, etc. as one for which we Wish ^a recognition machine to give an identifying re- sponse or alarm. It is not intended as ^a military term.

However, since the sensory apparatus of the MARK I severely limits ground resolution and has no gray scale, the use of actual air photos was prohibited and could only be approximated.

PRELIMINARY EXPERIMENTS WITH SIMPLE FORMS

Early in the program there were run, several experiments of a type called "one-vsmany," in which the machine is trained to respond to some one shape, called "target." Following such training, it is tested on targets and many other shapes as well. It is expected to respond to the target shape whenever it appears (in any position or size permitted by the experiment) and to ignore all other shapes.

The first experiments of this sort used alphabet letters, with X as the target, and *A, E, K, N,* 0, *R, S, T, U* as non-targets, all of which were permitted to appear in any upright position at one standard magnification. Tests showed a perfect detection rate for *X's* but a high over-all false alarm rate from the non-targets, distributed as shown in Table I.

In the next experiments, a group of subject shapes was chosen which were intermediate in general character between alphabet letters and objects on the ground, as seen from the air. The result was an improvement in discrimination, evidenced by a drop in the falsealarm rate. The shapes were an X-shaped cross, a circle, a square, a triangle, an arbitrary hatchwork of lines, and a moderately irregular distribution of irregularly shaped dots. They are called, respectively, \times , \odot , \Box , \triangle , Lines, and Dots, and are pictured in Figure 1. The Lines and Dots patterns were large enough to more than fill the retina, and various portions of the patterns were arbitrarily chosen and presented at different times in anyone experiment. They were also used as distributed backgrounds in later experiments.

In one set of experiments, X was the target and \circ , \Box , \triangle , Lines, and Dots the nontargets while, in a sister set, \circ , was the target and X, \Box, Δ , Lines and Dots were the nontargets. The results are given in Table **II.**

Note in this table that a substantial part of

FREQUENCY OF "TARGET" RESPONSE

	EXPT. $40 \times$ AND 40.5								
			Detection	False- Alarm					
					Lines	Dots	Rate	Rate	
Perceptron Trained on \times	78/78	3/79	7/24	7/24	10/49	6/24	1.0		
Perceptron Trained on \bigcirc	13/78	79/79	17/24	8/24	12/49	7/24	1.0	.29	

TABLE II

the false alarms for the O-trained perceptron came from \Box , which is the figure most resembling O. Similarly, the X-trained machine experienced the most difficulty with "Lines" in whose random hatchwork are many incidental \times 's. The false-alarm rate is considerably improved over that obtained with alphabet letters, but is still uncomfortably high. Now, it may be that, in some practical circumstances, detection may be of such importance that one would be willing to pay a fairly high price, in false-alarms, for extremely

FIG. 1A. Actual Patterns Projected onto Mark 1 retina.

FIG. lB. Pictures show how Projected Pattern in Fig. 1A "looked" to the Perceptron because of Quantization by the Low Resolution Retina.

630 PHOTOGRAMMETRIC ENGINEERING

TABLE III EXPT. 40.5 (No NEGATIVE TRAINING)

high detection reliability for bona fide targets. On the other hand, there are means, discussed below, for reducing false-alarms in machines of this type.

MULTIPLE LOOKS AND NEGATIVE TRAINING

In conventional operation, an output response is asked for and is given as soon as a stimulus image is fixated upon the retina of the MARK 1. On the other hand, people often make several fixations upon a pattern before making a final decision. To test the benefit of multiple fixation, stimulus patterns were projected in four successive, systematically related, positons on the retina, and an ultimate response was calculated for each set of four immediate responses. This technique is referred to by the term "4-fix", in Table III.

Some substantial reduction in false-alarm rate was obtained from 4-fix responses. Additional reduction should also be obtainable from negative training. In this case, the machine is taught not only to give positive

responses to targets, but to give no such responses to non-targets. In general, the class of non-target objects likely to be encountered by a photo interpretation machine is very large compared to the class of targets. Consequently, while it may be practical to train on a large fraction of all the possible target stimuli, this is not likely to be practical for non-targets. Training, no matter how extensive within reason, is not likely to cover more than a small sample of the possible nontargets. Therefore, an ability to generalize is even more important in the reduction of falsealarms than is actual target recognition.

In the next experiment, a perceptron trained to recognize X also received some negative training on a small sample of \circ , Lines, and Dots. When trained to recognize 0, it also received some negative training on X, Lines, and Dots. Neither perceptron received very thorough training on any of its non-target forms, nor did either perceptron receive any training on \Box or \triangle . The results, are shown in Table IV. They may be com-

FREQUENCY OF "TARGET" RESPONSE

TABLE V FROM EXPT. 40.5

pared directly with the result of the previous experiment "Experiment 40.5 (No negative training)." (Table III).

To observe the effect of partial negative training, the two tables should be compared row by row. One would expect to find some reduction in the false-alarms from just those stimulus forms for which negative training was given. But there appears to be some carry-over, or generalization, of the effect, not only to untrained *positions* of a partly negative-trained non-target shape, but also to other untrained non-target shapes. In fact, there appears to be even a slight (undesirable) reduction in *target* recognition. But a larger experiment is needed to determine whether it
is statistically significant Furthermore statistically significant. Furthermore, several cycles of target, non-target training can reduce this effect if it is significant.

DOUBLE STIMULI AND BACKGROUND

The experimental results reported above concern the effects of single and multiple

fixation, with or without negative training, on the detection of single "target" shapes in any upright position, and the rejection of single, isolated, "non-targets." The same experiments which produced those results, also yielded results on various combination stimuli, obtained by projecting any two of X, φ, \Box, Δ , Lines, and Dots simultaneously. Each member of any stimulus pair was placed on the machine's field of view in any upright position, randomly selected, and independent of the other member of the pair. The description of training already given still applies: the MARK I was split into two independent half-size perceptrons, one of which was trained positive on \times alone, one on \circ alone. As before, the experiments were divided into two sets which were identical except that, in one, some negative training was given on a small sample of single non-targets. The stimuli were then presented in pairs and the results are given in Tables V and VI.

These results, like those in Tables **III** and IV show an advantage for 4-fix over 1 -fix

		Stimuli									False-	
		XO	\times	$X \triangle$	$\frac{\times}{Line}$	$\frac{\times}{\textit{Dot}}$	O ₁	$\circ \triangle$	\circ Line	\circ	Detection Rate	Alarm Rate
Perceptron	$1-Fix$	6/12	10/12	$10/12$ 5/12 8/12			4/12		$3/12$ $1/12$ $2/12$.65	.21
Trained Pos. on \times	$4-Fix$	4/6	5/6	6/6	1/6	2/6	0/6	0/6	0/6	0/6	.60	θ
Perceptron	$1-Fix$	5/12	1/12	2/12	0/12		$2/12$ 9/12	4/12	$7/12$ 5/12		.50	.10
Trained Pos. on \bigcirc	$4-Fix$	3/6	1/6	0/6	0/6	0/6	6/6	6/6	3/6	5/6	.77	.04

TABLE VI FROM EXPT. 40.6

FREQUENCY OF "TARGET" RESPONSE

631

632 PHOTOGRAMMETRIC ENGINEERING

FIG. 2. Photographs show how several of the Combination Stimuli "looked" to the Perceptron, because of Quantization of the Projected Image by the Low-Resolution Retina. Combination Patterns were produced by projecting any independently, to new positions.

responses. The lowered detection rate for true targets is also plainly evident. Experience which was gained later indicates that this undesirable drop might have been substantially eliminated by continued, alternating, positive and negative training. Inherently difficult tasks, such as mixed stimuli, normally require several cycles of training to reach a practical plateau in performance, whereas the early experiments reported in this paper provided only one.

It is now important to find out, in the near future, by repeated cycles of positive training on targets and negative training on a sample of non-targets, whether one can decrease the false alarm frequency for all non-targets without a comparable reduction in target detection rate.

"REALISTIC" STIMULI

The pair presentation experiments represent the first baby step toward the recognition of well-defined target shapes in the presence of other shapes and backgrounds. The next series of experiments, in which clutter played a prominent part, represent a somewhat closer approach to realistic conditions, within the limited capabilities of the sensory apparatus available for the MARK I Perceptron.

Because of resolving power and gray-scale

limitations of its input device, only a very small fraction of the total information present in any area of an ordinary photograph could actually enter the MARK 1. This degradation at the input tends to reduce a good photograph to an unintelligible hodgepodge of black spots. As a consequence. ordinary aerial photographs could not be used directly. On the other hand. since a lack of ground. or gray-scale. resolution is not an inherent characteristic of perceptrons, it was permissible, for these preliminary tests. to compensate for the input inadequacies by artificially restoring the pattern continuity which would have been destroyed by direct photographic input to the MARK 1.

Figure 3 shows typical realistic "target" shapes in real photographs. It can be seen that extremely low-resolution and. especially, quantizing to black-and-white, would leave little. if anything. to be recognized.

Figures 4 and 5 show stimulus patterns derived from tracings of aerial photographs. After the outlines were filled in to produce solid silhouettes, they were copied on to 35 m m. slides. In the following experiments. background, consisting of some arbitrarily chosen subsection of the "Lines" or "Dots" slides. was independently added. through the use of a second slide projector and a beam splitting mirror. When subjected to the

633

FIG. 3.

FIG. 4B. Pictures showing how (A) in Figure 4A and combinations of (A) with (B) and with (C) appear after quantization by the Mark I Perceptron Retina. These are the visual data which actually enters the machine and upon which its decisions are based.

Alc IN REVETMENT

EMPTY REVETMENT SHED-TRENCH COMPLEX

FIG. 5A. Pictures showing original slides used in the Revetment Experiment.

A/C IN REVETMENT **EMPTY REVETMENT** SHED-TRENCH COMPLEX

FIG. SB. Pictures showing how the Mark 1 Retina Quantizes thc [mages in Figure SA. Quantized images, indicating what information actually enters the Perceptron, are obtained from the Sensory System Monitor. A Bank of $20'' \times 20''$ Neon Indicator Lamps, wired in parallel with the Sensory Units.

degradation imposed by the MARK I input apparatus, the lines or dots images lost much of their specialized character and became a suitable simulation of a splotchy, partly organized, rather dense, background pattern.

An experiment to illustrate directly the detection of targets in clutter, consisted of training the MARK 1* to a ship pattern and then testing its ability to detect this pattern

* Actually, only a half-size MARK I.

on a high noise background, such as would be furnished by specular sea clutter. As before, the stimulus was allowed only one rotational orientation, but no restrictions were applied to its location in the field of view. Responses were recorded for 12 different arbitrary sampling positions. The scores are shown in Table VII.

In another experiment, the MARK I was trained to detect revetments containing aircraft. Its ability to respond correctly to these

TABLE VII FREQUENCY OF "TARGET" RESPONSE

TABLE VIII FREQUENCY OF "TARGET" RESPONSE

A/C in	4/C	Revetment	Shed-Trench	Detection	False-Alarm	
Revetment	Alone	Alone	Complex	Rate	Rate	
11/12	2/12	0/12	0/12		.06	

and to reject empty revetments, aircraft not in revetments, and other airfield structures was tested as follows: Positive training was given for the "aircraft in revetment" stimulus of Figure 5. Some negative training was given for a few samples of "empty revetment." No training of either kind was given on the "shedand-personnel-trench complex" nor on the "aircraft without revetment." In tests following training, each of these stimuli was allowed to fall upright on any location in the field of view and the recorded responses are shown in Table VIII.

LATTICE RECOGNITION

Many patterns of practical interest are not characterized so much by a distinguishing outline shape as by some other identifying properties. One of these is sometimes periodicity. While periodic structures can occur naturally, quite commonly they are manmade objects such as tank farms, orchards, or urban grids. To a human observer, it may be the regular periodicity as much as any other property, which arrests the attention when these are encountered in an aerial photograph.

An exploratory experiment was completed on the MARK I to test whether some representation of this property could be trained into and used by a simple random perceptron. If it could be, the simple perceptron, which is normally very sensitive to size and outline characteristics, could be expected to recognize small portions of the original training lattice, regardless of its outline or size, or any reasonable portion of some other lattice whose spacing constant was somewhat close to that of the training lattice.

A regularly spaced dot pattern typical of a tank farm, was projected to fill the entire field of view of the MARK 1. At this magnification, each dot was larger than but was centered on one sensory unit of the 20×20 unit retina. Several training presentations of this image were made, shifting it sideways little-by-little between presentations until it had moved a total of one lattice unit where, again, each dot fell exactly upon one sensory unit. Similar training was given in the vertical direction. Also, and this is important, rotational variation of up to 90° was included in this experiment; by means of a dove prism.

After training, the perceptron was tested on various fragmentary portions of the original lattice, in several magnifications, and in several lateral, vertical, and rotational positions, with and without clutter backgrounds. Table IX shows the results obtained from 20 tests of each stimulus. These 20 were a random sampling of the permissible locations and rotations.

As one would expect, the fidelity of response falls off as the image magnification is made more and more different from that used in training. If the changes in magnification had been carried further to, say, 0.5 and 2.0, one would expect some rise in this performance because the sub-harmonic and harmonic patterns thus produced would so closely resemble the original pattern. The effect

t Because of the high symmetry properties of the pattern.

636 PHOTOGRAMMETRIC ENGINEERING

TABLE IX

FREQUENCY OF "TARGET" RESPONSE

would, of course, be blurred because of the translations and rotations which were permitted. But, if one point of the lattice were anchored on one particular sensory field point, and no changes in rotation were permitted as the magnification were decreased, the image at O.SX would be indistinguishable (to the perceptron) from a 5×5 section of the lattice at a magnification of LOX. There were no data for a 5×5 section at 1.0X but 80% of the responses to a 4×6 section at this magnification were correct.

CONCLUSIONS AND FORECASTS

There has now been completed the de scription of those preliminary experiments most directly related to potential applications in photo interpretation. They have demon strated photo recognition potential in a relatively simple device, and have directed attention to problems which the next phase of effort should attack.

One of the most evident needs, and perhaps the easiest to provide, is a moderate degree of gray-scale resolution in the input apparatus. When this is furnished, actual photographic images may be presented to the recognition device without serious loss of their characteristic continuity. On the other hand, preservation of the continuity in the image brings one face-to-face with the very difficult but necessary task of deciding exactly how continuity can be sensed and utilized in the

1. recognition of objects not characterized

- by fixed outline shapes, and,
- 2. segregation of target images from each other and from their backgrounds.

These two uses are major problem classes which appear to embrace most of the outstanding challenges for the mechanization of photo interpretation. For the recognition of patterns on some basis other than rigid shape, one obviously must rely on attributes which are not particularly related to shape (such as color, texture, periodicity, straightness, and certain topological relations). For segregation of patterns, one is led to consider means for detecting compactness and coherence of forms and to consider a hierarchical logical structure within the machine to deal with the hierarchical organization of ground patterns, in which properties and forms are often, if not usually, the elemental parts of larger patterns.

It is fortunate that these considerations are compatible with what we believe can be engineered into advanced machines whose ancestry can be traced to a simple perceptron of the MARK I type. Of the many specific problems which must be surmounted to reach these goals, three are here suggested for immediate attention.

1. A *search for pattern properties* which are moderately primitive but significantly more sophisticated or abstract than those on which a simple perceptron performs its classification, and the design of mechanisms for incorporation in a recognition machine to sense and utilize these properties in the identification of ground objects.

- II. A *two-pronged attack* on the sensing and utilization of continuity, beginning at the most elementary level, and aimed primarily at the recognition[†] of boundaries or edges.
	- (a) The most general case appears to be the easiest: the sensing of threadlike continua without regard for the regularity of their direction. It seems likely that the ability to sense threads is basic to the desired ability to recognize closed boundaries which at least define the existence and location of an object even if they do not identify it. Furthermore, the class of figures which may be called ribbons, including roads, rivers, and some beaches, consist primarily of two threads whose separation tends to remain constant. Therefore, one of the early goals to be sought is the development of mechanisms for recognizing rivers, roads, and other ribbon-like forms having generally unpredictable length and regularity of direction.
	- (b) The more difficult case appears to be the recognition of *regularity* in thread-like continua, particularly straightness. Any practical machine is likely to have a sensory mosaic similar to a retina but substantially coarser than that of a human eye. A perfectly straight boundary projected onto a man's retina in a fixed arbitrary position would excite a pattern of receptor cells whose boundary is irregular. Yet, somehow, the processing apparatus which we carry in our heads obtains a sense of straightness from such an irregular representation of a "truly" straight line and, to a remarkable degree, detects irregularities which are much smaller than the size of its own retinal mosaic elements. This

t As opposed to merely emphasizing or isolating them in a two-dimensional representation.

ability is usually stated in terms of vernier acuity.

A significant part of the problem of straight-line detection in machines is apparently in the problem of providing vernier acuity of an order greater than the retinal resolution, and an explanation of the mechanism in people may very well lead to a solution of the problem for machines, or vice versa. Because it appears to be so closely related to the topic of continuity, mechanisms for improvement of apparent vernier acuity are included as an immediate research goal for the continuity study.

Ill. The third problem area deserving immediate attention is simultaneous classification at several levels. It is related to topic II in the sense that both bear upon the so-called figureground problem of perceiving, as separate entities, the elementary objects of an ensemble of objects. Multiple-level classification processes may be thought of as those which may give several responses to a complex pattern; the responses must form an orderable set corresponding to the orderable levels of organization within a complex which consists of patterns, subpatterns, sub-sub-patterns, etc. It is not difficult to provide an elementary ability of this sort in primitive perceptrons and some of their foreseeable descendants, and it is intended to give special attention to this possibility in the immediate future.

REFERENCES

- 1. Rosenblatt, F., "A Theory of Statistical Sepa-rability in Cognitive Systems," C.A.L. Report
- No. VG-1196-G-1, January 1958, 263 pp.
ASTIA No. AD204076.
2. Murray, A. E., "A Review of the Perceptron
Program," Proceedings of the National Elec-
tronic Conference, Vol. XV (1959).
3. Murray, A. E., "Perceptron Applicab
- PICS). CAL. Report No. VE-1446 G-1,
- November, 1960, ⁵⁸ pp. 4. Platt, John R., "How We See Straight Lines," *Scientific American,* 202, No.6, June 1960.