INVARIANT RECOGNITION OF GEOMETRIC SHAPES

by

J. Alan Steppe

Eugene Wong

Memorandum No. ERL-M237

4 December 1967

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Invariant Recognition of Geometric Shapes*

J. Alan Steppe and Eugene Wong

Department of Electrical Engineering and Computer Sciences
and Electronics Research Laboratory
University of California
Berkeley, California

1. INTRODUCTION

In using geometrical figures to convey information, we make little use of the position or orientation of the figures. Rather, the information is contained in their "shapes," which might be defined as those properties of a figure which are invariant under rigid-body motions. An occasional counterexample, such as the numerals 6 and 9 in certain type fonts , only serves to remind us how rare these counterexamples are. It is not surprising, therefore, that most pattern classification problems for geometrical figures involve classes each of which is closed under rigid-body motions. Seldom, if ever, are we required to distinguish a figure from a congruent copy of itself. This paper is concerned with the problem of classifying geometric figures without regard to their positions and orientations. We want to establish procedures which recognize squares of a fixed size in any position and orientation as belonging to the same class.

In formulating pattern recognition problems it is often convenient

*The research reported herein was supported by the U.S. Army Research Office -- Durham under Grant DA-ARO-D-31-124-G776 and Contract DAHCO4-67-C-0046.

to distinguish two separate operations in the recognition procedure, namely, measurement and decision. The problem of measurement is to obtain from the sample to be classified a finite set of numbers which are indicative of the class to which the sample belongs. This operation has the dual function of data conversion and data reduction. The operation of decision consists of assigning the sample to one of the prescribed classes on the basis of the measured values. The bulk of the pattern recognition literature is devoted to the problem of designing effective decision procedures. We shall have something to say about both of these operations. It is intuitively clear that for recognition of geometric shapes, we should begin with measurement which are invariant under rigid-body motions. There are at least two reasons for this. First, such measurements coupled with any decision procedure automatically give rise to invariant recognition. Congruent figures will automatically be assigned to the same class. A second advantage of starting with invariant measurements is efficiency. Positional and orientational information is not relevant to the classification problem and should not be measured.

As was pointed out by Pitts and McCulloch [1], one way of obtaining invariant measurements is to average over all rigid-body motions. This idea of averaging is also basic to the suggestion of Novikoff [2] that the techniques of integral geometry be used in pattern recognition. The basic principle underlying Novikoff's suggestion can be illustrated by a simple example. Suppose that the geometric figures to be recognized are simple closed curves. Suppose we throw straight lives "randomly" at the figures and measure the lengths of their intersections. If "random" is taken to mean that no position or orientation is given a preference, then the average of these intersections will converge to a number which is independent of the position and orientation of the unknown figure. Hence, the average constitutes an invariant measurement. This example also serves to illustrate some of the difficulties associated with the averaging approach. First, to obtain a good approximation to the limiting average may require a large number of lines to be thrown, which is clearly undesirable. Secondly, and more importantly, the average, while invariant with respect to rigid-body motions, may not depend much on shape either. For example, the average of the lengths of intersection between a simple closed curve and random lines is proportional to the area enclosed by the curve. Clearly, this measurement is useful only in discriminating among figures with different areas. Furthermore, since only approximations involving a finite number of lines can be obtained in practice, the figures to be distinguished should have very different enclosed areas, in order for the average length of intersection to be a useful measurement. Invariance is a necessary property for measurements, but not a sufficient property. The most important attribute of a measurement is, after all, its power in discriminating between classes.

While the approach suggested by Novikoff has been developed in somewhat greater detail by Ball [3] and Tenery' [4], there appears to have been no study specifically addressed to the two difficulties mentioned earlier; namely, how to keep the number of trial measurements small and how to choose measurements which are not only invariant with respect to motions but also discriminating with respect to shape. This paper summarizes come preliminary results of a study to re-examine the appropriateness of integro-geometric techniques in the recognition of geometric shapes, with special attention being given to the difficulties mentioned earlier. There are two distinct features in our approach which should be pointed out at the very outset. First, we have essentially abandoned the idea of averaging, in favor of treating the observations as "random variables." Secondly, sequential analysis is used and plays an important role in keeping the number of observations small. The preliminary results from a computer study are very encouraging indeed. In assessing these results, however, we must keep in mind certain assumptions. First, these results pertain to noise-free recognition problems. Each class contains only figures congruent to each other, and not deformed figures. In practice, of course, deformations are unavoidable, and they will change the results somewhat. Secondly, these results pertain only to two-category classification problems. The basic procedure extends readily to the multi-category situation, but the more optimistic conclusions may not. Finally, as will be seen, the invariance that we actually achieve is with respect to only those motions which keep the figure within a prescribed retina. The larger the retina, the larger will be the needed number of observations for a given level of performance.

2. INTEGRAL GEOMETRY IN THE PLANE

Let R^2 denote the Euclidean plane, on which we can choose an origin and Cartesian coordinates (x, y). Euclidean distance between two points (x, y) and (x_o, y_o) is given by

(1) $$d[(x, y), (x_o, y_o)] = [(x-x_o)^2 + (y-y_o)^2]^{1/2}$$

By a rigid-body motion g we mean a mapping of R^2 onto R^2 which is one of the following types:

(a) g is a translation $t_{(\xi, \eta)}$, $(\xi, \eta) \in R^2$,

(2) $$t_{(\xi, \eta)}(x, y) = (x + \xi, y + \eta)$$

(b) g is a proper rotation τ_θ , $0 \le \theta < 2\pi$,

(3) $$\tau_\theta (x, y) = (x \cos \theta - y \sin \theta, y \cos \theta + x \sin \theta)$$

(c) g is a succession of translations and rotations.

The collection G of all rigid-body motions is a group. Every rigid-body motion is equivalent to a translation followed by a rotation. This correspondence $g = \tau_\theta t_{(\xi, \eta)}$ provides a natural coordinate system $\{(\xi, \eta, \theta); (\xi, \eta) \in R^2, \theta \in [0, 2\pi)\}$ for G. Hence, a subset A of G corresponds to a subset $\bar{A}$ of $R^2 \times [0, 2\pi)$. Consider those subsets A for which the integral

$$\mu(A) = \int_{\bar{A}} d\xi \, d\eta \, d\theta \tag{4}$$

is well-defined. The function $\mu(A)$ has the important invariance property

$$\mu(g_o A) = \mu(A) \quad \text{for every } g_o \text{ in } G \tag{5}$$

where $g_o A = \{g_o g;\ g \in A\}$. Furthermore, subject to some mild additional conditions, every function $F(A)$ with this invariance property is a constant multiple of $\mu(A)$. The function $\mu(A)$ is known as an invariant measure (or Haar measure) for the group of rigid-body motions.

As illustrated by the example in the introduction, the measurements that we shall consider are all generated by throwing lines, points, arcs and other geometric objects randomly against the unknown figure which is to be classified. The notion of "randomness" is to be defined in such a way as to be consistent with invariance with respect to rigid-body motions. The way to do this is through the use of the invariant measure. Let ω_o denote a fixed set of points in R^2. For example, ω_o may be the horizontal axis $\{(x, 0),\ -\infty < x < \infty\}$, or the origin $(0, 0)$. Let M be the collection of all possible copies of ω_o under rigid-body motions, i.e. $M = \{g\omega_o,\ g \in G\}$. It turns out that the invariant measure μ of G induces automatically an invariant measure μ_M on M. That is, $\mu_M(\cdot)$ is a non-negative σ-additive set function defined for a class of subsets of M such that

$$\mu_M(g_o B) = \mu_M(B) \quad \text{for } g_o \in G$$

By choosing a coordinate system for M, μ_M can be expressed similarly to (4). For example, let M be the set of all infinite straight lines. Every such line can be obtained from the horizontal axis $\omega_o = \{(x, 0), -\infty < x < \infty\}$ by a vertical displacement $t_{(o, y)}$, $-\infty < y < \infty$, followed by a rotation τ_θ, $0 \le \theta < \pi$. Thus, $(y, \theta) \in (-\infty, \infty) \times [0, \pi)$ provides a coordinate system for M. A subset B of M corresponds to a subset $\overline{B}$ of $(-\infty, \infty) \times [0, \pi)$, and up to a positive multiplicative constant $\mu_M(\cdot)$ is given uniquely by

$$(6) \qquad \mu_M(B) = \int_{\overline{B}} dy\, d\theta$$

Roughly speaking, integral geometry deals with integration with respect to the invariant measure μ_M for various choices of M.

Let C denote a fixed geometric figure in the plane. Let $f(\omega, C)$ be a real-valued function which depends on C and on elements ω of a given M in such a way that f is invariant when ω and C are simultaneously acted upon by a rigid-body motion. Then, the integral

$$(7) \qquad I(C) = \int_M f(\omega, C)\, \mu_M(d\omega)$$

is invariant when C is acted upon by rigid-body motions. This is the average of f. If M is the set of all straight lines, C is a simple closed curve, and $f(\omega, C)$ is the length of intersection between C and ω, then I(C) is proportional to the area enclosed by C. References 2 and 3 contain a number of examples where the integral I(C) can be explicitly computed. More examples can be found in the book by Santalo [5]. In practice, unless I(C) corresponds to a simple quantity such as area or perimeter, I(C) can only be <u>measured</u> by generating elements of M according to the invariant measure and average the results. Strictly speaking, only a modification of this can be implemented. Instead of all straight lines, we consider only those that intersect a prescribed retina, and modify the invariant measure accordingly. This will be explained in more detail subsequently. In most cases the number of elements of M that have to be generated may be quite large for a satisfactory approximation to I(C).

The process of estimating I(C) by observing f for a sequence of randomly generated elements from M is analogous to the process of estimating the mean of a distribution by the sample mean of a sequence of sample random variables. The correspondence becomes exact, if we convert the unnormed measure μ_M into a probability measure by the introduction of a retina. Let R, called the retina, be a closed and bounded region in the plane which contains the figure C. Let M_o be the subset of those elements of M that intersect R. For a subset A of M, we define

$$(8) \qquad \rho(A) = \frac{\mu_M(A \cap M_o)}{\mu_M(M_o)}$$

The function $\rho(A)$ is defined whenever $\mu_M(A)$ is defined and is a probability measure. If we are choosing elements from M according to the ρ measure, then elements that do not intersect the retina do not get chosen. Of course, ρ is no longer a true invariant measure. For a given A in M_o, $\rho(A)$ is invariant only for those motions which do not take A out of M_o. The process of estimating I(C) can now be stated precisely as follows: For the function $f(\omega, C)$, let

$$(9) \qquad P(x) = \rho(\{\omega; f(\omega, C) \leq x\})$$

be its probability distribution function. The mean of the distribution is equal to

$$(10) \qquad m = \int_{-\infty}^{\infty} x \, d\, P(x) = \frac{1}{\mu_M(M_o)} \int_{M_o} f(\omega, C)\, \mu_M(d\omega)$$

which is not quite I(C) but closely related to it. If, for example, $f(\omega, C) = 0$ whenever ω does not intersect the retina R, then $m = \frac{1}{\mu_M(M_o)} I(C)$. Now, the sequence of observations on f generated by choosing elements out of M corresponds to a sequence of identically distributed random variables: $f_1(\omega, C)$, $f_2(\omega, C)$, ..., each being distributed with distribution function P(x). The random variables should preferably be chosen to be independent as well. The problem is reduced to one of determining whether the sample mean

$$m_N(\omega) = \frac{1}{N} \sum_{k=1}^{N} f_k(\omega, C)$$

converges to m.

3. SEQUENTIAL ANALYSIS

From the point of view of pattern recognition there is no reason why we should restrict our attention to sample average. Instead, we should seek the best procedure for deciding the class to which C belongs, using as data a sequence of identically distributed samples. The situation is this: The random variable $f(\omega, C)$ has a distribution function $P(x, i)$ which depends on i, the class to which C belongs. Let $f_1, f_2, \ldots, f_N$ be a sequence of random variables independently and identically distributed according to the distribution function of f. Upon observing a realization of the sequence $f_1, \ldots, f_N$, we want to determine the class number i. Thus posed, the problem becomes a standard problem in statistical classification. In fact, we have here a much nicer statistical problem than is usually encountered in pattern recognition. Normally, the measurements in a pattern recognition problem are neither independent nor identically distributed. It is also difficult to think of another pattern recognition situation where it is natural not to fix the number of measurements, namely, N, a priori. The setting here is perfect for the application of sequential analysis [6].

The basic ideas of sequential analysis can be outlined as follows. Our observation is represented by a sequence of independent and identically distributed random variables $f_1, f_2, \ldots$.. Given a set of observed values, we are to decide something about the common distribution of $f_1, f_2, \ldots$.. The number of observations in a sequential test is not a priori limited.

After n observations we can either go on observing or terminate the observations and make the necessary decision concerning the distribution of $\{f_i\}$. Unlike the fixed-sample case, we now have an additional degree of freedom, namely, the flexibility of determining when to stop observing. In sequential analysis one usually tries to minimize the average number of observations consistent with a given level of performance. To be specific, suppose f_1, f_2, ... can be distributed according to one of two distribution functions, with density functions $p_1(x)$ and $p_2(x)$ respectively. Now, there are two types of error that can be made; deciding that the density function is $p_1(x)$ when it is in fact $p_2(x)$, and conversely. Suppose we require that the probability of error for these two types of error not to exceed α and β respectively, and design a decision procedure which minimizes the average number of observations that are required to achieve this level of performance. The resulting decision procedure is the celebrated sequential ratio test of Wald [6]. First, two positive numbers A and B, with $B > A$, are computed for the given values of α and β. Suppose f_1, f_2, ... are the observed values, then for each $n \geq 1$ we compute the likelihood ratio

$$L_n(f_1, \dots, f_n) = \prod_{k=1}^{n} \frac{p_1(f_k)}{p_2(f_k)} \tag{11}$$

and compare L_n against A and B. As long as L_n lies between A and B, sampling continues. As soon as L_n exceeds B or falls below A, the sampling stops, and we decide in favor of $p_1(x)$ of $L_n > B$, and conversely if $L_n < A$. Clearly, the number of observations depends on the actual values observed, and is thus a random variable.

The sequential procedure can be easily adapted for the pattern recognition problem that we have been discussing. In order to obtain a better assessment for the entire approach to invariant recognition, some numerical computations have been conducted, and more experiments are underway. Preliminary results are summarized in the next section.

4. EXPERIMENTAL RESULTS

All the experiments reported in this section have the following features in common:

(a) M is the set of all infinite straight lines.

(b) The geometric figures to be recognized are all simple closed figures.

(c) $f(\omega, C)$ is the length of intersection between the straight line ω and the figure C.

(d) In each experiment only two classes are involved, and a version of the sequential ratio test is used.

In each experiment two simple closed line figures representing the two classes are chosen. Roughly speaking, the sizes of figures are chosen so as to render the underlying classification problem as difficult as possible. For a fixed retina R, the probability distribution functions $P_1(x)$ $P_2(x)$ corresponding to the two figures are then computed for $f(\omega, C)$ using (8) and (9). We next compute the average number of samples needed for a given pair α and β, probabilities of the two types of error. The main variables of each experiment are α, β and the size of the retina R.

The role of the retina is an important one, because it pertains to the question of invariance. The recognition procedure is invariant only with respect to those motions which keep C within the retina. Naturally, we can expect that the larger the retina, the larger the required sample size. This comes about because for a large retina there is a large probability that the observed intersection is zero, which is not an informative piece of data. Provided that the retina is convex, the average sample size depends only on the perimenter of the retina and not its shape. However, the specific motions that we can tolerate obviously depends on the shape of the retina.

The distribution functions P(x) that we encounter in these experiments can always be written as

$$
\begin{aligned}
(12) \qquad P(x) &= P_o + (1 - P_o)\int_0^x p(x')dx' \qquad && x \geq 0 \\
&= 0 && x < 0
\end{aligned}
$$

where $P_o \geq 0$ is the probability of the length of intersection being zero, and p(x) is a probability density function. As was alluded to earlier, the principal effect of increasing the size of the retina, is to increase P_o. Let $P_i(x)$, $i = 1, 2$ be the distribution functions corresponding to the two classes, and let P_o^i, $p_i(x)$, $i = 1, 2$ be defined in terms of $P_i(x)$ as in (12). Define for $x \geq 0$ the function L(x) by

$$(13) \qquad L(x) = \frac{P_o^1}{P_o^2}, \qquad x = 0$$

$$= \frac{p_1(x)}{p_2(x)}, \qquad x > 0$$

For a sequence of observation $x_1, x_2, \ldots,$ let

$$(14) \qquad L_n(x_1, \ldots, x_n) = \prod_{i=1}^{n} L(x_i)$$

The sequential test consists of computing L_n and comparing it against A and B as was explained earlier. If the retina is large, then $\frac{P_o^1}{P_o^2}$ is approximately 1, and both P_o^1 and P_o^2 are nearly 1. This means that there is a large probability that L_{n+1} and L_n are approximately equal. Obviously, this situation would require a large number of samples to reach a decision.

In our computer study, a set of 15 simple basic figures were chosen. This generates 105 pairwise recognition problems. Of these, 5 pairs are presented in the accompanying figures. They are arranged in an increasing order of difficulty, ranging from the trivial problem of separating a circle from a block U, to the rather difficult problem of separating two nearly identical pentagons. In each of these cases, the probabilities of error of the two types have been set to be equal, and two values of this probability for each retina ratio are presented. Retina ratio refers to ratio of perimeters, and retina ratio = 1 is for any retina which encloses either shape in its standard position, and has the same perimeter as the smallest circle which has this property. The average sample size depends on the true class of the figure being tested, and both cases are presented.

We feel that these results are very encouraging. For an example of medium difficulty, consider Case 3: circle vs. square. For a retina ratio of 1 and an error probability of .2 per cent, no more than 5 samples are required. For a retina ratio of 2 and the same error probability, the average sample size is 8 or 15 depending on the true class of the observed figure. These results indicate that the procedure proposed in this paper is well within the realm of practicability. Case 5 involves a problem of discriminating two nearly identical pentagons. This is by far the most difficult case among the 105 that were studied. We feel that the difficulty for this case is only in part due to the close similarity of the two shapes, but perhaps more importantly, it is due to the fact that line intersections are ill-suited for discriminating between these two shapes. For a given set of shapes the problem of determining the best manifold M of random objects and the best invariant function $f(\omega, C)$ to be used is undoubtedly the most important open problem in the application of integral geometry to pattern recognition. We feel that the techniques outlined in this paper provide an adequate vehicle for the comparison of two integro-geometric measurements. However, the problem of determining the "best" measurement is not only unsolved, but is yet to be properly posed. We are continuing our study in this direction.

REFERENCES

1. W. Pitts and W. S. McCulloch, "How We Perceive Universals: The Perception of Auditory and Visual Forms," Bull. Math. Biophysics, Vol. 9, pp. 127-147, 1947.

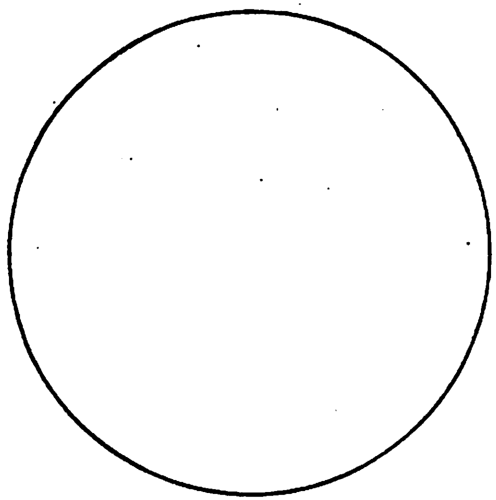
2. A. B. J. Novikoff, "Integral Geometry as a Tool in Pattern Perception," Principles of Self-Organization, Von Foerster and Zopf, ed., Pergamon Press, pp. 347-368, 1962.

3. G. H. Ball, "An Application of Integral Geometry to Pattern Recognition," Stanford Research Inst., Report for Office of Naval Res., Contract No. 3438/00, 1962.

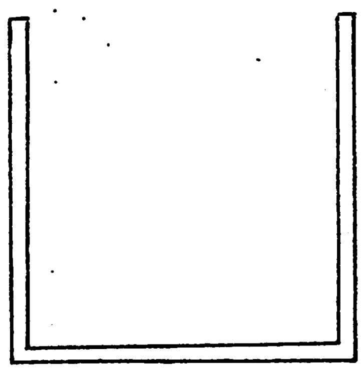
4. G. Tenery, "Pattern Recognition Function of Integral Geometry," IEEE Trans. on Military Electronics, Vol MIL-7, pp. 196-199, 1963.

5. L. A. Santalo, Introduction to Integral Geometry, Hermann and Co., Paris, 1953.

6. A. Wald, Sequential Analysis, John Wiley and Sons, Inc., New York, 1947.



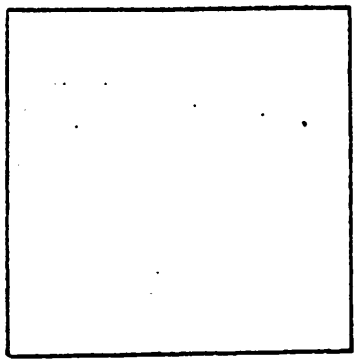
CLASS 1 CIRCLE



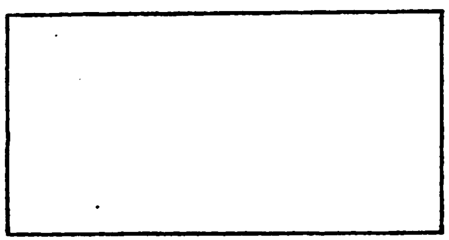
CLASS 2 BLOCK U

RETINA RATIO	ERROR RATES	AVERAGE SAMPLE NUMBER OBSERVING 1	AVERAGE SAMPLE NUMBER OBSERVING 2
1	0.01	0.616	1.00
	0.002	0.646	1.37
2	0.01	1.25	2.30
	0.002	1.31	3.17
10	0.01	6.24	11.5
	0.002	6.53	15.8

Fig. 1



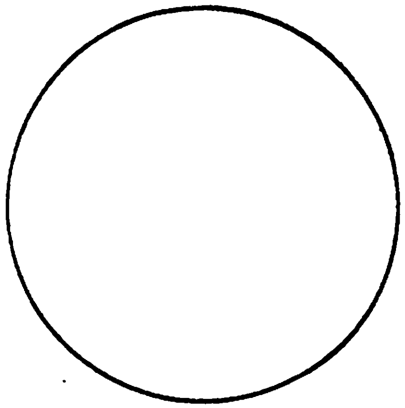
CLASS 1 SQUARE



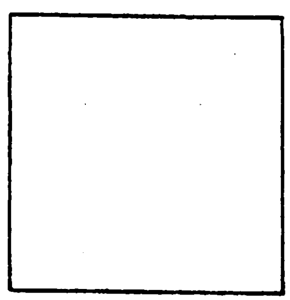
CLASS 2 2:1 RECTANGLE

RETINA RATIO	ERROR RATES	AVERAGE SAMPLE NUMBER OBSERVING 1	AVERAGE SAMPLE NUMBER OBSERVING 2
1	0.01	5.79	7.97
	0.002	7.96	11.0
2	0.01	11.7	16.2
	0.002	16.1	22.3
10	0.01	58.6	81.3
	0.002	80.6	112.

Fig. 2



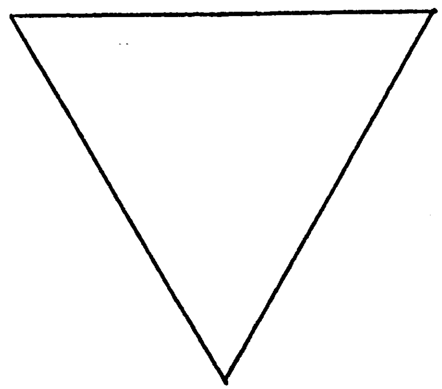
CLASS 1 CIRCLE



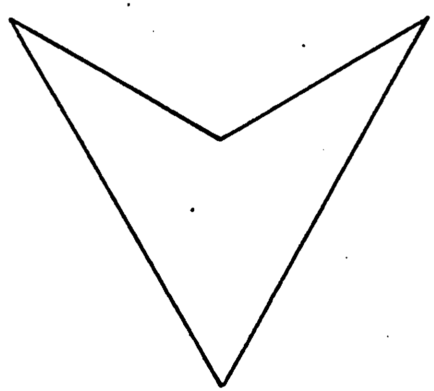
CLASS 2 SQUARE

RETINA RATIO	ERROR RATES	AVERAGE SAMPLE NUMBER OBSERVING 1	AVERAGE SAMPLE NUMBER OBSERVING 2
1	0.01	2.72	3.19
	0.002	3.73	4.39
2	0.01	5.76	10.9
	0.002	7.92	14.9
10	0.01	28.9	54.7
	0.002	39.7	75.1

Fig. 3



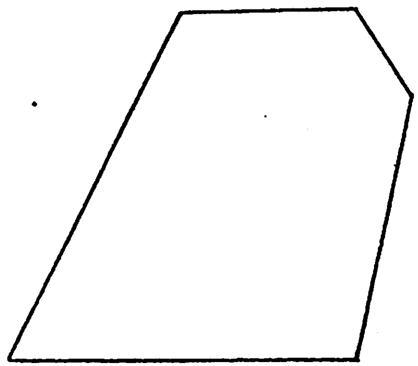
CLASS 1 TRIANGLE



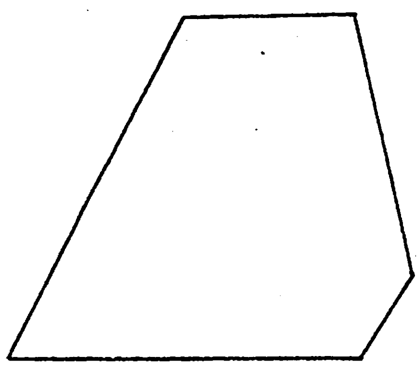
CLASS 2 V

RETINA RATIO	ERROR RATES	AVERAGE SAMPLE NUMBER OBSERVING 1	AVERAGE SAMPLE NUMBER OBSERVING 2
1	0.01	24.6	29.8
	0.002	33.8	40.9
2	0.01	49.2	59.5
	0.002	67.6	81.8
10	0.01	246.	298.
	0.002	338.	409.

Fig. 4



CLASS 1 PENTAGON 1



CLASS 2 PENTAGON 2

RETINA RATIO	ERROR RATES	AVERAGE SAMPLE NUMBER OBSERVING 1	AVERAGE SAMPLE NUMBER OBSERVING 2
1	0.01	261.	271.
	0.002	359.	373.
2	0.01	522.	543.
	0.002	717.	746.
10	0.01	2610.	2710.
	0.002	3590.	3730.

Fig. 5