# VISION
## HUMAN AND ELECTRONIC

Albert Rose

*David Sarnoff Research Center*
*RCA*
*Princeton, New Jersey*

PLENUM PRESS • NEW YORK-LONDON

*CHAPTER 1*

# THE VISUAL PROCESS

## 1.1. Introduction

It would be difficult to find a more cogent confrontation between physics and biology than in the visual process. Nature was faced from the beginning with the hard fact that light consists of a finite number of bits of energy, called "photons" or "quanta." Whatever visual information was to be distilled out of the surrounding world was circumscribed by the profound constraints imposed by the discrete nature of light.

Throughout the millions of years over which life in its manifold forms evolved, survival was the dominant motif. Unless the prey could detect its predator in ample time, life terminated abruptly. Visual detection was not the only means of detection, but it was a major one. And visual detection had still to function in twilight and even in starlight when the stream of photons dwindled to an occasional patter of drops of energy. It was, indeed, a matter of life and death that each photon be husbanded and assembled to trace out the best possible image of impending disaster. Little short of a photon counter would suffice.

There is ample evidence among the primitive forms of life that nature mastered the art of counting photons at an early age. If it were only a question of utilizing the incident energy of the photons, we can, in fact, predate animal life and point to the highly efficient solar battery developed in plants by way of photosynthesis. But the counting of photons entails not only the efficient absorption of photons but also the highly sophisticated process of amplification. The energy of a photon is sufficient to disturb only a single atom or molecule. With this energy alone, the information that a photon had been absorbed could not be transmitted beyond the point of absorption, let alone to some central nervous system. A nerve pulse involves the motion of at least some millions of atoms or ions. Hence, the energy of the absorbed photon must be multiplied or amplified over a millionfold before it can give rise to a nerve pulse. The ingenious amplifier that nature devised remains an unsolved puzzle. The great variety of amplifiers that man has devised for the same purpose is in large part the content of this monograph.

The quantum character of light is a hard constraint. Nature could, in a physical sense, do no more and, in a survival sense, do no less than devise a photon counter. Once having the photon counter, there were secondary choices as to how the information was to be handled.

The incoming photons, for example, could be accumulated for a long time to generate an image of high quality or for short times to give a rapid series of low-quality images. The long-time accumulation would mean that moving objects would be blurred. Moreover, the animal itself would have to slow down its movements so that it did not joggle its camera (or visual system) during the course of an exposure. At the other extreme, a radically shortened time of exposure would yield images of such poor quality or so impoverished of information content as to be of little value in guiding the animal's response. The compromise was set, at least for the human system, where one might expect, namely, at an exposure time matched to the reaction time of the human system as a whole. The reaction time is the sum of the transit time of nerve pulses from eye to brain and back to an appropriate extremity plus the time required to overcome the physical inertia of that extremity. Overall, the reaction time is in the order of a tenth of a second, as is the exposure time of the eye.

The choice of exposure time is readily understandable. So also is the choice of spectral response. The latter peaks near the peak of the sun's radiation—and even shifts at twilight toward the blue in order to match the shifting spectral content of the light scattered from the "blue" sky. There are a host of other choices perhaps not so obvious and perhaps offering the opportunity for reading the past through the shape of the present. These have to do with the size of lens opening, the focal length of the lens, the red and the blue cutoffs of the visible range, the density of retinal elements which puts a ceiling on image quality, the multiplicity of color vision, the programed interconnections of the optic nerve fibers, and, finally, even the number and location of eyes. We will return to this subject later and point out some of the adaptations of the optical parameters to the life habits of a number of animals. In the meantime, we note the contrast between the primary character of the photon-counting problem—it is singular, ultimate, and essential—and the wide-ranging secondary character of the ways in which the photon counting was adapted to the life habits of particular animals.

## 1.2. Quantum Limitations on the Visual Process

The absolute measure of the performance of a visual system is the ratio of information transmitted by the system to the information incident on the system and contained in the incident light flux. It is necessary, then, that we have a quantitative measure of the information conveyed by a finite number of photons. We derive such a measure in a series of steps designed to emphasize three aspects of the quantum limitations on the visual process. The first aspect has to do with the overall finite number of photons; the second aspect is their random distribution in time and space; and the third aspect is the problem of guarding against false alarms, that is, spurious visual patterns that may arise from the random character of the photon distribution and not from the original scene itself. Each of these aspects exacts an increasing toll in terms of the number of photons required to transmit an elementary bit of information.

### *1.2.1. Discreteness of Light Quanta*

We imagine first that we have a black canvas and that we wish to paint a picture (Fig. 1.1) on the canvas depicting a white wall on which is located a single black spot. This is the simplest

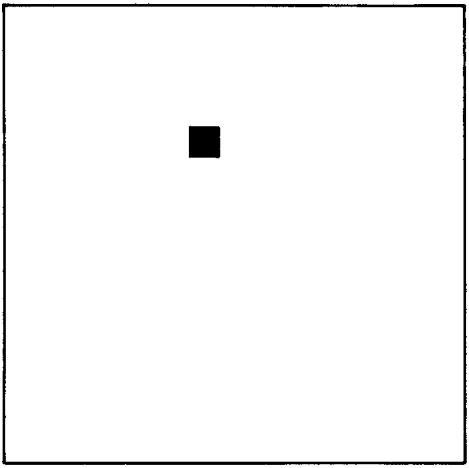

Fig. 1.1. Test pattern consisting of a single black spot on a white background.

of pictures in that we wish to indicate only the presence of the black spot, and not its structure, on the otherwise white wall. Furthermore, the method of painting will be constrained to be that of stippling. We can paint an array of small white dots, all of the same size but of varying spacing. Each white dot will correspond to the visual effect of a photon in a generalized visual system.

We suppose that the size of the black spot is such that the canvas can accommodate a total of $N$ of these close-packed spots. The single black spot, then, defines the size of a picture element whose area is a fraction $N^{-1}$ of the canvas.

At this point, we ask what would be the smallest number of white dots required to portray the presence of one black spot on a uniformly white wall? If we are allowed to space the white dots uniformly, then, clearly, $N - 1$ white dots are both necessary and sufficient to complete the canvas. The single missing white dot locates the presence of a single black spot (Fig. 1.2).

The next step in sophistication of our painting will be to locate a single gray spot as well as to portray its shade of grayness. We assume that the reflectivity of the spot is 99% of that of the

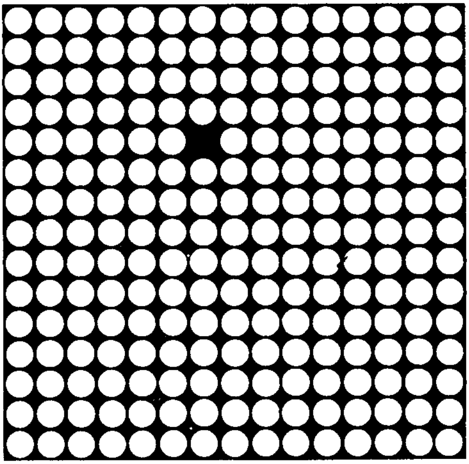

Fig. 1.2. Reproduction of Fig. 1.1 using an ordered array of "photons."

white wall and again ask for the smallest number of white dots required to convey this information. The answer, obviously, is $100N - 1$ dots. Each picture element will have exactly 100 dots stippled in, with the exception of the one picture element containing the gray spot. The latter will have 99 dots to indicate that its brightness is 99 % of that of the surrounding wall.

While all of the above is embarrassingly elementary, the argument does stress the high cost in photons required to portray small elements of low contrast. For example, the number of picture elements $N$ required for well-resolved images often lies in the range of $10^6$–$10^7$. Hence, we would need some $10^8$–$10^9$ photons to delineate the location and brightness of the gray spot.* That is,

* While a single gray spot on a white wall may in one sense appear to transmit only a single item of information, the total picture transmits as much information as any more complex pattern. The picture contains the information that there is *no* gray spot on any of the other $N - 1$ picture elements. Hence, the brightness of each picture element is a discrete independent item of information even when the picture is a complete "blank."

we would need these $100N$ photons providing they could be arranged in a precise array of 100 photons per picture element. But nature does not work in so orderly a fashion. Photons arrive at random times and places and give rise to a fundamental graininess in any image, a graininess that tends to obscure the detection of fine detail and faint contrasts. The result is a considerable increase in the number of photons required to delineate the fine detail of images.

### *1.2.2. Random Character of Photon Distributions*

Natural incoherent light is emitted by some form of electronic excitation, as from an excited atom. The average lifetime of the excited state is a well-defined, calculable, and observable parameter. On the other hand, it is a fundamental property of the quantum-mechanical nature of such an excitation that the photon can be emitted at any time during the average life of the excited state. More definitely, the *probability* of emitting a photon at any time $t$ and in a time interval $\Delta t$ is given by $\exp(-t/\tau)\,\Delta t/\tau$, where $\tau$ is the average lifetime of the excited state. What is significant for our purposes, is that the emission of photons is a stochastic process.

If we carry out the experiment of illuminating a small area with a "constant" incandescent light source and count the number of photons that strike the area in a given time $\Delta t$, we will obtain a series of numbers $n_1, n_2, \ldots$ corresponding to the actual number of photons that arrived at the first interval $\Delta t$, the second interval $\Delta t$, and so on. We have put quotation marks around the word "constant" because the experiment which we are performing is one way of determining whether the source is indeed constant. The fact that the numbers of the series $n_1, n_2, \ldots$ do not all have the same value may in the broadest sense cast doubt upon the constancy of the source. Yet, no matter how carefully we design the source, it will turn out that there is an irreducible spread to these numbers. That spread is the consequence of the stochastic or random nature of the process of emitting photons.

In particular, it will be true that if we observe that the average number of photons arriving at the test area is $n_0$, we will also find that the numbers $n_1, n_2, \ldots$ are distributed around $n_0$ in such a fashion that the average value of $(n_i - n_0)^2$ will also be

$n_0$. The average value of $(n_i - n_0)^2$ is called the mean-squared deviation from the mean. The square root of this quantity, or $\langle (n_i - n_0)^2 \rangle^{1/2}$ is called the *root mean squared deviation* from the mean and is abbreviated as *rms deviation.*

We introduce here also a terminology that will be used frequently in this monograph. A *signal* is defined as the *average* number of photons falling on a test element. The *noise* is the *rms deviation* from this number. In the example cited above, $n_0$ is the value of the signal and $n_0^{1/2}$ is the value of the noise. The signal-to-noise ratio is then also $n_0^{1/2}$. The term signal will also be used, as will be clear from the context, to mean the *difference* between the average numbers of photons falling on a given test element and on surrounding test elements of the same size. This meaning is used, for example, to define the signal appropriate to a low-contrast test-element on a uniform surround.

We return to the black canvas on which we wish to paint a small gray spot by use of a stippling of white dots, each representing a photon. In our first estimate, using uniformly spaced dots, we arrived at the need for 100 dots per picture element (a picture element was defined as the area of the gray spot to be portrayed) in order to portray a single gray spot having 99 % of the brightness of the surrounding canvas. If we now recognize the random character of the photon distribution, we will find that the actual numbers of photons falling on various picture-element areas are distributed around the average number 100 such that the rms deviation is $(100)^{1/2}$ or 10.* At this point, we have a signal to be detected which is 1 % of the surrounding average brightness and we are faced with a noise fluctuation in these picture element areas which is 10 % of the average brightness. (The signal to be detected in this example is the difference between the number of photons in the test element and the average number of photons in equal area elements of the surround.) In brief, the signal-to-noise ratio is 0.1, and far less than the value of unity which is frequently taken as the threshold for visibility of a signal against a noisy background.

* The rms deviation will be the same whether we look at the numbers of photons falling on a given area in successive equal time intervals or at the numbers of photons falling on many equal areas in a single time interval.

Our first requirement, then, is to increase the density of photons on the canvas so that the fluctuation or noise level does not exceed the signal to be detected. Since the signal represents a 1 % deviation from the surround, we require that the noise level (or rms deviation) also not exceed 1 %. This is achieved by having an average of $10^4$ photons falling on each picture element. The rms deviation will then be the square root of $10^4$, or $10^2$. And the ratio of this random deviation to the average will be $10^{-2}$, or 1 %.

In summary, at this point, the number of photons required to portray a single spot was increased by a factor of 100 in going from a black spot to a gray spot having only 1 % contrast with the surround; and the number was increased again by a factor of 100 in going from an ordered array of photons to a random array. The latter factor insured that the signal to be detected was equal to the rms deviations occurring as a result of the fundamentally random character of the photons. There is yet another factor to be introduced to guard against false alarms, that is, the mistaking of any particular random fluctuation for the real signal to be detected.

### *1.2.3. False Alarms*

It is frequently stated or implied in discussions of electronic systems that the threshold of detectability of a signal occurs when the signal is equal to the noise. This is a somewhat misleading statement. For example, suppose that we are monitoring the level of an electrical current to detect significant changes of 1 % or more. Suppose, also, that the noise level (rms deviation) is 1 % of the average current. If we make $N$ successive observations of the current, we will find that the current will depart from its average value by 1 % or more in almost half of the observations even in the absence of any "real" or deliberately imposed disturbance of the current. We must put quotes on the term "real" because the fluctuations resulting from the noise are just as "real" as a deliberately imposed disturbance—it is only that the source of the disturbance is different. The significant part is that almost half of our observations will tell us that the current has departed from its average value by more than 1 % whether or not there has been any deliberate disturbance. It is in this sense that half the observations will be false alarms. In order to guard against false alarms, the "real" signal to be detected must exceed the level of noise by some appropriate factor. The factor can readily be approximated by knowing the statistical distribution of noise fluctuations as well as the number of observations which should statistically be free from false alarms.

Figure 1.3 shows the distribution of noise fluctuations around the mean value of a parameter. The ordinate is the probability density and the abscissa $k$ is plotted in units of the rms deviation. The second abscissa scale, $n$, is a particular numerical example for which the average number of photons is 900. The rms deviation is then 30. The total area under the curve using the $k$ abscissa scale is unity. The area under the curve between $k = 1$ and $k = 2$, for example, is 0.13 and represents the probability that an observation will lie in the range between 1 and 2 rms deviations above the mean. In the numerical example, it is the probability that an observation will lie between 930 and 960 photons. Similarly, the

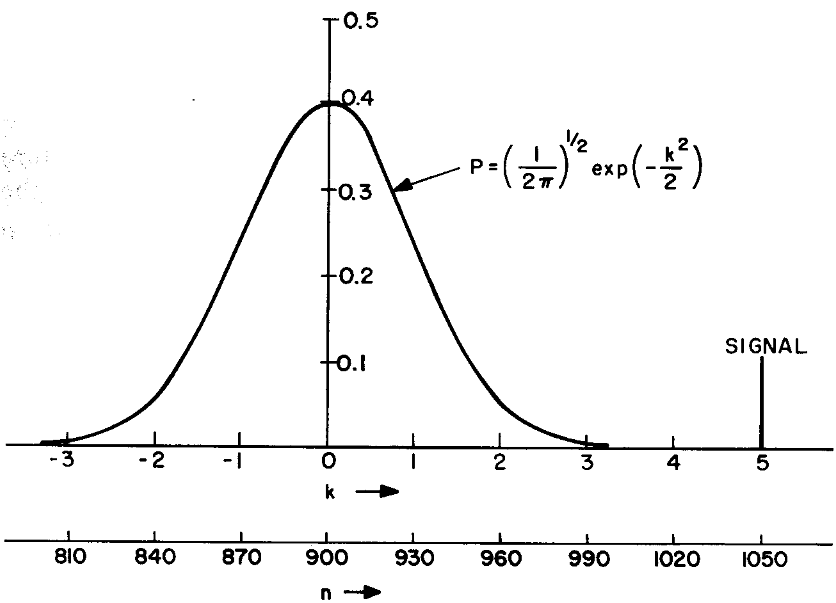


Fig. 1.3. Probability distribution of a noisy quantity about its mean value.

area under the curve to the right of $k = 2$ is 0.023 and represents the probability that an observation will exceed 2 rms deviation units above the mean. In the numerical example, it is the probability that an observation will exceed 960 photons.

Table 1.1. gives the probability that noise fluctuations will exceed the mean value of the background by 1, 2, 3, etc. units of the rms value of the noise. The probability for noise fluctuations occurring on either side of the mean is just twice the probabilities listed.

With the aid of Table 1.1 we can now specify how large a signal is required in order to avoid false alarms. The signal in this case is the difference in average brightness between the test spot and the background. We suppose, as is common, that the picture has $10^5$ picture elements, each of the size or area occupied by the test spot. We have then $10^5$ opportunities to generate a false alarm. And, if our purpose is to reduce the number of false alarms to below unity, we will need, according to Table 1.1, a signal whose amplitude is 4–5 times larger than the rms noise. We call this value of $k$ the *threshold signal-to-noise ratio.* It is such a value of signal for which we are reasonably confident of not mistaking a noise fluctuation for the real signal. Note that at $k = 6$ the probability of detecting a false alarm is already far smaller than is needed. In the other direction, $k = 3$ would only guard against false alarms for a picture having less than $10^3$ picture elements. Hence $k = 5$ is a reasonable approximation to the threshold signal-to-noise ratio.

**Table 1.1**
**Values for the Probability of Exceeding Various Values of $k$**

| $k$ | Probability of exceeding $k$ |
|---|---|
| 1 | 0.15 |
| 2 | 0.023 |
| 3 | $1.3 \times 10^{-3}$ |
| 4 | $3 \times 10^{-5}$ |
| 5 | $3 \times 10^{-7}$ |
| 6 | $2 \times 10^{-9}$ |

We choose $k = 5$ rather than $k = 4$ for the following reason. In the above argument, we assumed a very noisy background and a well-defined signal. However, the signal itself has nearly the same noise or spread as the background.* This means, for example, that if in Fig. 1.3 we located the mean value of the signal at $k = 4$, we would find that the signal appeared half the time below $k = 4$, and half the time above. If, on the other hand, we locate the signal at $k = 5$, we will find that only 0.15 of the time will the signal appear below $k = 4$. It will exceed $k = 4$, on the average, 0.85 of the time and be judged a real signal. Hence, a margin of about one unit of $k$ above the nominal value needed to avoid false alarms is sufficient to give a reasonable reliability to our observations.

At this point, we compute the increase in photon density required to satisfy the criterion $k = 5$, as compared with the criterion $k = 1$ used in the previous section. We begin with the condition $k = 1$ for which the signal is equal to the rms deviation of the noise. In particular, let the signal and the rms deviation each be 1% of the background brightness. As we increase the photon density, the signal remains constant when measured as a percentage of the background brightness. The rms deviation of the noise, however, decreases. Since the ratio of the rms deviation to the average background brightness varies as $n_0^{1/2}/n_0 = 1/n_0^{1/2}$, where $n_0$ is the average density of photons in the background, it will be necessary to increase $n_0$ by a factor of $k^2$ ($= 25$) in order to decrease the ratio $1/n_0^{1/2}$ by $k$ ($=5$).

In sum, the density of photons required varies as $k^2$. And, for the value $k = 5$, the density of photons must be increased 25-fold relative to the density computed in the previous section for $k = 1$. In the previous section, the number of photons was computed to be $10^4 N$, where $N$ was the number of picture elements. Hence, to guard against false alarms, this number must be increased to $2.5 \times 10^5 N$.

We can now write, in general, the expression for the total number of photons required to detect a contrast $C$ where $C$ is a

* Black or very dark signal elements on a white background must, of course, be excepted.

measure of the signal as a fraction of the background brightness, that is, $C \equiv \Delta B/B$ and $0 \leqq C \leqq 1$ ($C = 1$ means 100% contrast and $C = 0.01$ means 1% contrast).

$$\text{Total number of photons} = N \frac{1}{C^2} k^2 \tag{1.1}$$

Here, $N$ is the total number of picture elements and reflects the discreteness of the photons. The factor $1/C^2$ is a consequence of the contrast $C$ and the random character of photon distributions; the factor $k^2$ reflects both the random character of the photon distribution and the need to avoid false alarms.

## 1.3. A Summary Experiment

Almost all of the conclusions of the previous three sections can be read off by inspection of Fig. 1.4. In Fig. 1.4a there is depicted an area uniformly illuminated by a low density of photons. Each photon was made visible on a television screen as a discrete white dot by using a high-gain photomultiplier. We note in Fig. 1.4a the discrete character of photons, their random distribution, and their consequent noisiness which gives rise to false alarms. A simple inspection of Fig. 1.4a reveals black areas or spots which we could, in the absence of other information, readily identify as "real" black spots in the original picture. In fact, no such deliberate pattern of black spots was introduced into the making of Fig. 1.4a. The black spots are a consequence of the statistical fluctuations in the distribution of photons.

Figure 1.4b is a "real" test pattern of black spots* which we wish to detect under the low illumination represented by Fig. 1.4a. To do so we simply superimposed the positive transparencies of Figs. 1.4a and 1.4b to obtain Fig. 1.4c. Note that the four larger black spots of Fig. 1.4b are readily visible in Fig. 1.4c. The remaining smaller black spots of Fig. 1.4b are undetectable in Fig. 1.4c. They are lost in the noise. Note also that the four larger spots that are visible in Fig. 1.4c *appear* to terminate in a fifth black spot forming the apex of a triangle. The fifth black spot, however,

* The thin black bar will be referred to in Section 1.5.

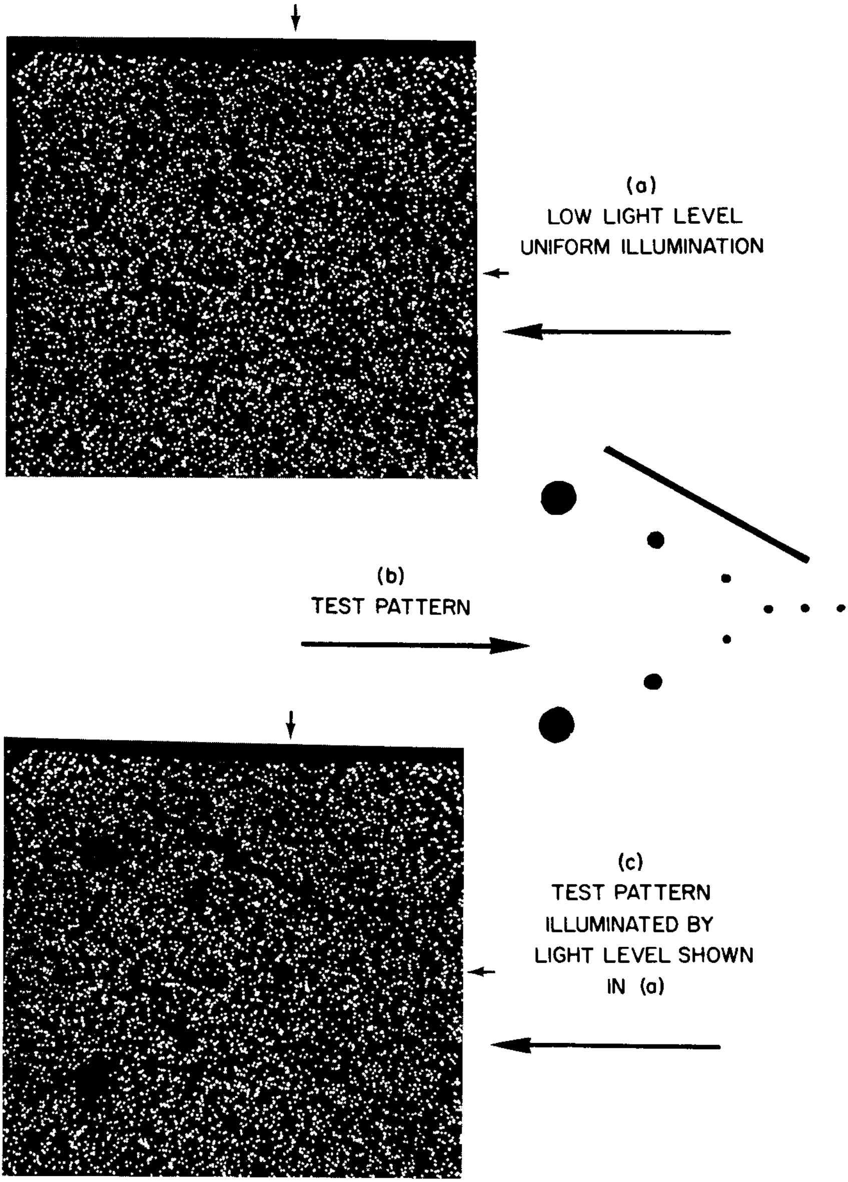


Fig. 1.4. Demonstration of the limitations imposed by the quantum nature of light on its ability to transmit information.

is one of the largest statistically generated black spots *already* present in the "uniform" illumination of Fig. 1.4a. For convenience, this black spot is located in Figs. 1.4a and 1.4c by small coordinate arrows on the edges of the pictures.

The presence of this statistically generated black spot in Figs. 1.4a and 1.4c means that any "real" black spot must be larger than it in order to be reliably judged to be "real." If we take this statistically generated black spot as a starting point, we find that it occupies about 1/500 of the area of the picture. Also, since there are some 4500 dots in the picture, the average number of dots in the area of this spot is 9. The signal-to-noise ratio is then $\sqrt{9} = 3$. From Table 1.1, we note that a signal-to-noise ratio of 3 would yield false alarms only one in a thousand times. Hence, since there are 500 picture elements having the size of the black spot, and since the probability of obtaining such a black spot by statistical fluctuation is one part in a thousand, we are at the threshold of reliable visibility, of "real" signals or "real" black spots. The "real" black spots need only be somewhat larger than the one we are considering. In particular if the signal-to-noise ratio of the "real" spot were between 3 and 4, we would have sufficient confidence in its reality. Note, for example, that the signal-to-noise ratio of the largest black spots in Fig. 1.4c is approximately 5, since they each obscure about 25 white dots. These large black spots are clearly well above the threshold of reliable visibility and tend to confirm that our estimate of threshold signal-to-noise ratio should lie between 3 and 4.

The density of photons in Fig. 1.4 is at the extreme low end of densities that we normally encounter. The low density was deliberately chosen to illustrate the three major properties of photon distributions: discreteness, random distribution, and false alarms. In the range of the higher densities normally encountered, the number of picture elements is likely to be of the order of $10^6$ rather than the $10^3$ calculated for Fig. 1.4. Under these conditions, the threshold signal-to-noise ratio also must be increased to values between 4 and 5 in order to guard against the appearance of false alarms. From Table 1.1., at $k = 5$, the probability of false alarms is only $3 \times 10^{-7}$. It decreases rapidly at $k = 6$ to $2 \times 10^{-9}$. A television picture, for example, has some $10^5$ picture elements and would call for a value of $k$ between 4 and 5.

Figure 1.4a serves another purpose. It emphasizes how completely unrealistic it would be to use the customary criterion for visibility, namely, a signal-to-noise ratio of unity. The operational meaning of this criterion is that if a first person removed one of the dots in Fig. 1.4a, a second person could determine which dot had been removed. The removal of one dot is the equivalent of having a black spot in a test pattern that obscures, on the average, one dot or photon. The average signal is then one photon and the noise, which is the square root of the average, is also one photon, yielding a signal-to-noise ratio of unity. Simple inspection of Fig. 1.4 shows the virtual impossibility of detecting a single missing photon.

## 1.4. A Second Experiment

The experiment of Fig. 1.4 was confined to the visibility of black spots for which the contrast is, by definition, unity. If we choose to look for gray spots for which the contrast is less than unity, the size of the spot must be increased. According to Eq. (1.1),

$$\begin{aligned} \text{Total number of photons} &= N \frac{1}{C^2} k^2 \\ &= \frac{A}{d^2 C^2} k^2 \end{aligned} \tag{1.2}$$

where $d$ is the linear dimension of a picture element (that is, the test spot) and $A$ is the area of the picture. Equation (1.2) states that if we keep the total number of photons fixed (that is, maintain constant brightness), the diameter $d$ of a test spot which is just visible should vary inversely with its contrast $C$.

Figure 1.5 is a photograph of a test pattern made up of discs whose diameters decrease by a factor of 2 in progressing along a row, and whose contrasts decrease by a factor of 2 in progressing down a column. Along a 45° diagonal of Fig. 1.5, the product $dC$ is then constant. If we illuminate Fig. 1.5 with some intermediate light intensity, the boundary between the discernible and nondiscernible parts of Fig. 1.5 should be one of the 45° diagonals.

Figure 1.6 shows the results of a series of illuminations of Fig. 1.5. Figure 1.6 was obtained by photographing the kinescope of a television system when the pattern of Fig. 1.5 was illuminated

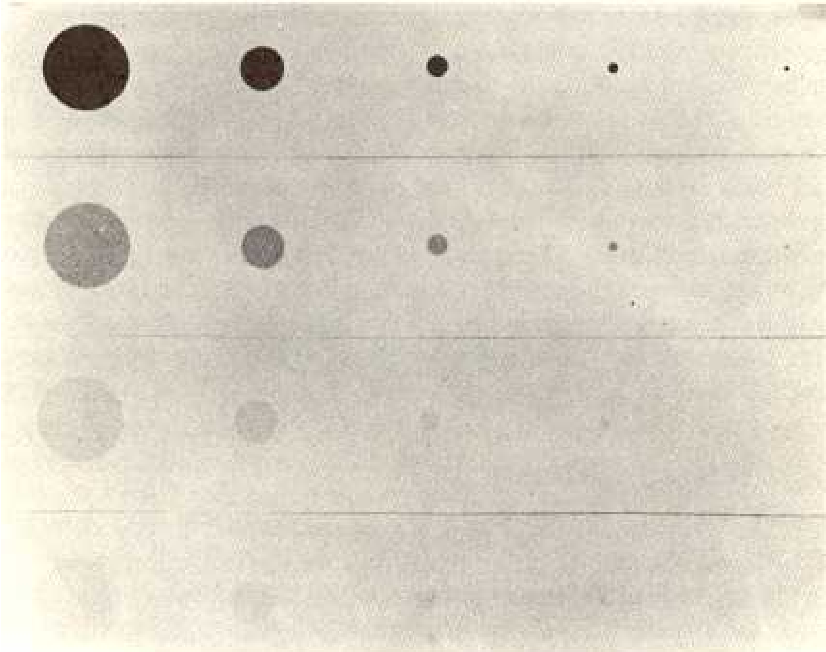

Fig. 1.5. Test pattern used to measure the resolving power of a system in terms of the size and contrast of single elements.

by a flying-spot scanner and the reflected light recorded by a high-gain photomultiplier.(R-1) Parenthetically, Fig. 1.6 was made some years prior to Fig. 1.4 and shows a certain variation in the intensities of the individual white dots, each being the trace or signal of an individual photon. The variation arises from the variation of gain of a photomultiplier depending upon where on the photocathode the photon strikes. In Fig. 1.4, a limiter was used to trim the white dots to nearly the same size.

Figure 1.6 shows clearly that the boundary between the discernable and the nondiscernable parts of Fig. 1.5 lies approximately along a 45° diagonal. Further, the boundary moves one step to the right, toward discs that decrease by a factor of 2 in diameter, for each factor of 4 increase in light intensity, as is to be expected from Eq.(1.2). The series of pictures in Fig. 1.6 was used in the early publication(R-1) to estimate a value for $k$, the threshold

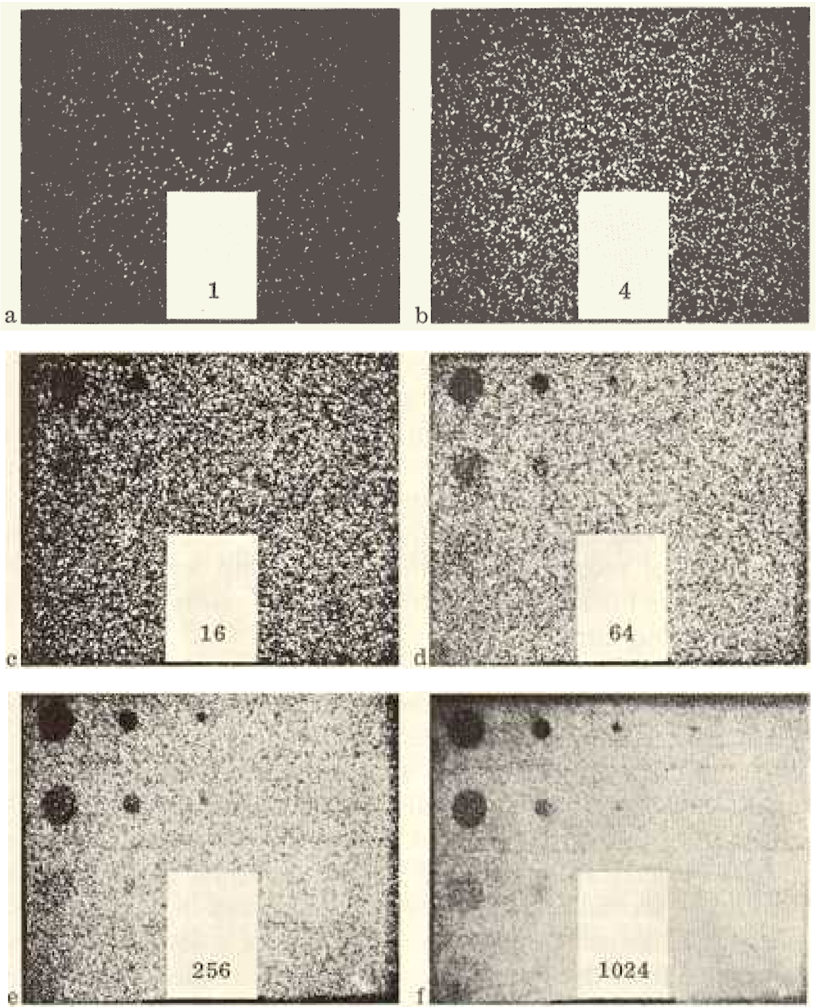


Fig. 1.6. Reproduction of Fig. 1.5 using a light-spot scanning arrangement in which the trace of single photons was made visible. The relative numbers of photons are indicated on each photograph.

signal-to-noise ratio, lying between 4. and 5. As we noted earlier (see Table 1.1), the value of $k$ should increase from 3 to 5 in going from very low densities of photons for which the number of picture elements is less than $10^3$ to high densities of photons for which the number of picture elements exceeds $10^5$.

## 1.5. Resolution, Signal-to-Noise Ratio, and Test Patterns

The term "picture element" has been used here as meaning the smallest area of a spot of a given contrast that can be resolved. The shape of the spot is not critical; it can be round, square, or even rectangular. The area is of primary significance in determining the signal-to-noise ratio which is based on the average number of photons falling on that area.

It is clear by inspection of Fig. 1.6 that the term picture element has a certain elastic significance. If we are looking for small black spots on a white surface, the number of discernable picture elements is larger than if we are looking for gray spots having a small contrast with the white surround. From Fig. 1.6 or from Eq. (1.2), the number of discernable picture elements is proportional to the square of their contrast. There are $10^4$ more picture elements having a contrast of 1 (black spots) than those having a contrast of 0.01, or 1%. A lack of recognition of this relationship was responsible for many years for an inflated estimate of the resolving capability of photographic film.

The quoted values for film resolution were (and also, frequently, still are) based on the smallest spacing of a set of *black and white bars* that could be resolved. It was not unusual, for example, to use a given film with a rating of say 2000 lines per picture and find that not even details having the dimensions of 400-line resolution could be resolved. The primary reason, of course was that the detail to be resolved was low contrast, and not black and white. A second reason is that the detail was in the form of a single picture element and not in the form of a set of bars. The use of bars, rather than the single spots of Fig. 1.5, leads to a gross overstatement of the resolving properties of a system. The bars yield a higher resolution because the estimate of signal-to-noise ratio tends to be based on the total area of the bars rather than on a single elemental area whose diameter is the width of one bar. In Fig. 1.4, for example, the smallest black spots are not visible in Fig. 1.4c. At the same time, the presence of a long bar, whose width is slightly less than that of one of the smallest black dots, is readily visible.

There was a period during the early history of television when members of the motion picture industry asked for television channel widths of the order of a hundred megacycles in order to transmit their films.[(K-1)] This was based on a black and white bar-pattern resolution rating for their film of 2000 lines. A casual inspection of current television pictures shows that viewing this same 2000-line film through the present 500-line (5 megacycle channel width) television standards frequently *degrades* the picture quality as compared with live studio pictures. That is, the signal-to-noise ratio of the film at 500 lines is frequently less than that which emerges from the studio camera transmitting live scenes.[(H-1)]

The misleading effect of using a test pattern of bars to define the resolution of a system is clearly brought out in Fig. 1.7. This is a series of pictures of ten pairs of black and white bars taken by Coltman.[(C-1)] The number under each picture gives the relative illumination of the test pattern. The illumination was sufficiently small (and the gain of the system sufficiently high) that each white dot represents a photon. It is clear that even in the first picture of the series, the one which has the lowest illumination, the presence of the bars can be detected, for example, by viewing the pattern somewhat edgewise, first from the side and then from the bottom. In this picture the density of photons is so low that if we defined a picture element as a square element whose length of side is the width of one bar, the average number of photons in this element would be significantly less than unity and in the neighborhood of $\frac{1}{4}$. The signal-to-noise ratio of such an element would then be $(\frac{1}{4})^{1/2}$, or 0.5. This is a misleading use of the term signal-to-noise ratio since it applies to an element which is small compared to what the eye is actually looking at or making use of in order to

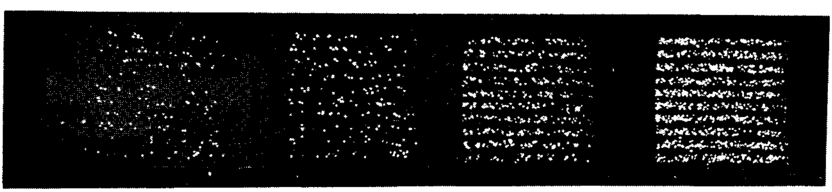


Fig. 1.7. Enhanced visibility of bar patterns as compared with dot patterns (Coltman).[(C-1)]

determine the presence of the bar patterns. The eye is making use of a large part of the pattern in order to achieve a signal-to-noise ratio well above unity. Single isolated elements whose dimension is the width of a bar would, of course, be completely undetectable in the first and also in the second picture of this series.

Consider also the first of the series of pictures in Fig. 1.6. This picture was recorded from the kinescope of a 500-line television system. If we were to estimate the signal-to-noise ratio of an element whose diameter is one television line width, the value, based on the density of photons, would be far less than unity. Since there are some few thousand photons in the entire picture, and a few hundred thousand television picture elements, the average number of photons per picture element is $10^{-2}$ and the signal-to-noise ratio referred to these elements is $10^{-1}$. It is clearly evident from Fig. 1.6 that elements of such small dimension and small signal-to-noise ratio are completely beyond the range of visibility. It is only by going to the area of the largest black spot that we attain a threshold visibility and a threshold signal-to-noise ratio of about 4.

In brief, the signal-to-noise ratio of what the eye is able to detect must be well in excess of unity. Signal-to-noise ratios less than unity, as reported by Coltman[(C-2)] or by Morgan[(M-1)] for the viewing of bar patterns, may have a certain utility as reference numbers, but they do not define the signal-to-noise ratio of what the eye actually apprehends.

Parenthetically, there is some ambiguity about the meaning of signal-to-noise ratios less than unity. For example, the signal-to-noise ratio of 0.1 cited above was associated with an average density of 0.01 photons per picture element. Hence, on the average, a viewer would see zero photons in the picture element 99 times out of 100 observations. A single observation does not, in general, give any information about the magnitude of signal-to-noise ratios less than unity.

Another frequent error arises in motion picture or television practice when the signal-to-noise ratio of a picture element in a *single* frame is associated with threshold visibility observations on the moving film. If the storage time of the eye were equal to the time for which one frame is viewed in motion pictures, the above association would be valid. As it is, the storage time of the eye is about 0.2 sec and the time for one frame (in a television system) is 0.03 sec. Hence, the eye is in effect looking at the superposition of some 7 successive frames and achieving a signal-to-noise ratio which is larger than that of a single frame by the factor $\sqrt{7}$. Anyone who has looked at a single frame of a motion picture is immediately aware that it is noisier than the visual impression gained from the moving film in normal projection.

## 1.6. An Absolute Scale of Performance

Using Eq. (1.2), an absolute scale of performance can be plotted against which the performance of any actual picture-seeing device or system can be measured. We choose a fixed value of 5 for $k$, the threshold signal-to-noise ratio, with the understanding that in the region of low light levels its value should be somewhat lower. There are no reliable measurements on the variation of $k$ with light level. Further we compute Eq. (1.2) for 1 cm$^2$ of image surface. Hence Eq. (1.2) can be written

$$\text{Number of photons/cm}^2 = n = \frac{25}{d^2C^2}$$

or

$$\text{Number of resolvable lines/cm} = \frac{1}{d} = \frac{Cn^{1/2}}{5} \tag{1.3}$$

The term "number of resolvable lines/cm" is used here to measure the diameter $d$ of the smallest resolvable isolated spot having a contrast $C$. It does not refer to the customary operation of viewing bar patterns with spacing $d$. The latter, as we have pointed out, are intrinsically more visible than single isolated spots. Figure 1.8 is a plot of number of lines/cm *versus* $C$ with the photon density $n$ as a parameter.

If we know, for example, that the image surface of some visual system has received an exposure of $10^{10}$ photons/cm$^2$, then, according to Fig. 1.8, we should be able to resolve black and white elements ($C = 1$) whose diameter is greater than $\frac{1}{2} \times 10^{-4}$ cm. At the same time, we should only be able to resolve elements, having a contrast of 0.01, whose diameter is greater than $\frac{1}{2} \times 10^{-2}$

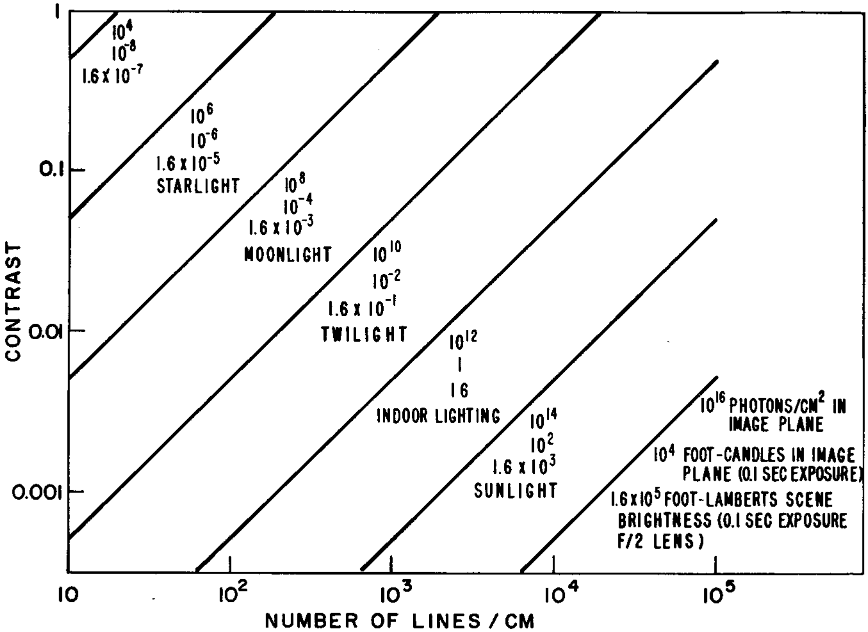


Fig. 1.8. Universal plot of the limiting size and contrast of single elements that can be transmitted at various photon densities.

cm. All of this would be true if, indeed, the visual system counted every incident photon. Under these conditions its quantum yield would be unity (a quantum *yield* of unity equals a quantum *efficiency* of 100%).

Now suppose that, in fact, the smallest black and white element that can be resolved by the visual system with $10^{10}$ photons/cm$^2$ on its image surface is only $\frac{1}{2} \times 10^{-3}$ cm rather than $\frac{1}{2} \times 10^{-4}$ cm. Then, according to Fig. 1.8, the effective performance is that of a system having a quantum yield of 0.01, or a quantum efficiency* of 1%.

The parameters of the curves in Fig. 1.8 are given primarily in photons/cm$^2$ in the image plane. For convenience, the photons/

* Quantum efficiency, here, has the same meaning as the phrase DQE (detective quantum efficiency) which was introduced by R. Clark Jones and which is used extensively in the literature.

cm$^2$ are converted also to foot-candles illumination in the image plane for an exposure time of 0.1 sec.* The conversion factor used is $10^{16}$ photons/sec per lumen of white light. A third equivalent shown on the curves is the corresponding scene illumination in foot-lamberts when an *f*/2 lens is used and an exposure time of 0.1 sec. Finally, the locations of various representative light levels are indicated from starlight to bright sunlight.

Figure 1.8 shows the performance of ideal noise-limited visual systems. It is likely that the resolving power of actual systems for small black test elements will be limited by lens errors, by diffraction, or by structure in the image plane. Similarly, the ability of actual visual systems to portray small contrasts in large areas may be limited by various sources of system noise such as nonuniformities in the recording medium. The result is that the performance curve for an actual system may not lie along a 45° line, but may be bowed such that the high-resolution and low-contrast ends show a lower performance than some intermediate part of the curve. Figure 1.6 shows evidence of this bowed type of cutoff. It will appear again in the peformance curves for human vision.

## 1.7. Geometric *versus* Noise Limitations to Performance

The preceding remarks are presented schematically in Fig. 1.9. The solid lines labeled "signal" give the amplitude response of an imaging system as a function of the number of lines/cm in a test pattern. These curves are a measure of the geometric limitations of the system such as diffraction, lens aberrations, and the finite size of the elements of the imaging surface. The curve labeled "rms noise" is the noise current that would be observed if the image were scanned by a series of apertures, each corresponding to a certain number of lines/cm. (The finer apertures give larger noise currents, increasing as the number of lines/cm, or the reciprocal width of the aperture. The reason is that even though the rms fluctuation within the aperture decreases as the aperture dimension, the linear velocity at which the aperture scans the image must increase as the reciprocal area, or as the square of its

* We assume here a lens transmission and a scene reflectivity of 100%.

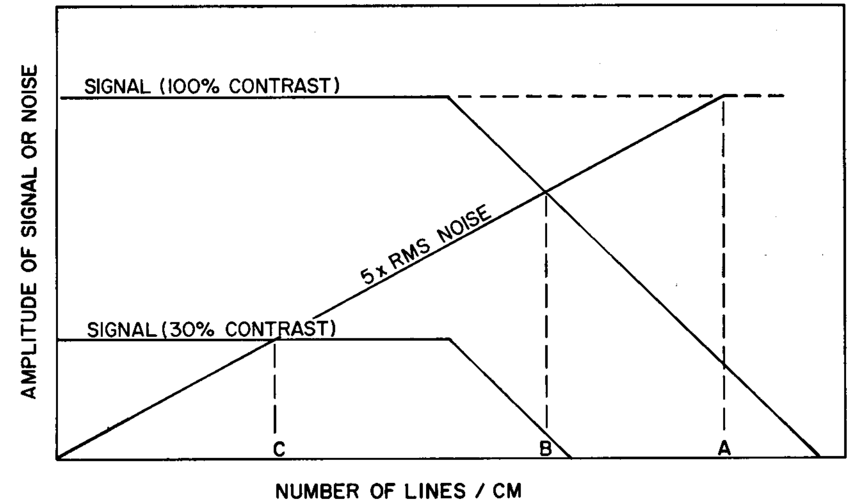


Fig. 1.9. Schematic comparison of geometric and noise limitations to resolution.

reciprocal dimension, in order to cover the entire image in a fixed time. The product of these two effects then yields a noise current increasing as the reciprocal of the aperture width. By the same arguments, the signal current is independent of aperture size.) The noise curve is plotted as 5 × rms noise in order that its intersection with the signal curve will yield directly the smallest resolvable elements.

If there were no geometric limitations to the resolution of the system, the smallest resolvable elements would lie at point *A*. Actually, the drop in the geometric resolution causes the cutoff to lie at a lower line number, at point *B*. This is true for picture elements having 100% contrast, that is, black and white. If we turn our attention to low-contrast elements [curve marked "signal (30% contrast)"] we find that the cutoff lies at point *C* where the geometric response is still unimpaired. The same is true for low-light scenes in general.

The point we wish to make is that even when geometric limitations on the resolution of a system begin to play a role, they affect selectively only the high-contrast parts of the picture. The visibility and signal-to-noise ratio of the low-contrast parts are likely still to be unaffected; and most of the information in an average picture is of low contrast.

Our emphasis throughout this monograph is on the limitations imposed by the finite number of photons rather than on the less fundamental limitations imposed by the finite geometric response of the system.

## 1.8. Beyond the Visible Spectrum

The arguments outlined in this chapter have been concerned primarily with radiation in the visible range of wavelengths (0.4–0.7 $\mu$m). Since the arguments have been couched in terms of numbers of photons, they apply equally well to any system that can detect ultraviolet radiation, x-rays, or gamma rays. Sturm and Morgan,[S-1] for example, have given an excellent discussion of the information transmitted by a finite number of x-ray photons. In the case of visible and higher energy radiations, one can extend the arguments down to almost arbitrarily low densities of photons since the thermal densities of these photons are vanishingly small.

One can also apply the arguments to the range of infrared radiation. Here, however, one must contend at ordinary temperatures with a significant density of thermally generated photons. It is as if the surround were never dark. Indeed, the ambient density of photons whose wavelength is in the neighborhood of 10 $\mu$m is comparable with that of bright sunlight, namely, about $10^{18}$ photons/cm$^2$-sec incident on or emitted from a surface. At 3 $\mu$m, the photon density is of the order of that for room light, or about 10 foot-lambert, and at 1 $\mu$m the density is equivalent to a visible ambient density of about $10^{-11}$ foot-lambert, that is, far below the absolute threshold for vision.

In general, the photon flux emitted by a blackbody at temperature $T$ is

$$\frac{\Delta\nu}{\lambda^2} \exp\left(-\frac{h\nu}{kT}\right) \text{photons} \cdot \text{cm}^{-2} \cdot \text{sec}^{-1}\,\text{sr}^{-1}$$

where $\Delta\nu$ is the range of optical frequencies in the neighborhood of the wavelength $\lambda$.

The visibility of objects in the infrared region is a complex function of the artificial illumination used, their self-luminous flux, their emissivities, and their temperature differences. We mention here only that the contrast of objects viewed by their own radiation and having the same emissivities is

$$\frac{h\nu}{kT}\frac{\Delta T}{T} \times 100\%$$

where $\Delta T$ is the temperature difference between an object and its surround. A temperature difference of 1 deg (centigrade) yields a contrast of about 10% at wavelengths of 1 $\mu$m and a contrast of about 1% at 10 $\mu$m.

## 1.9. Summary

The information content of a finite amount of light is limited by the finite number of photons, by the random character of their distribution, and by the need to avoid false alarms.

The signal-to-noise ratio of a test element in an image is defined as the ratio of the average number of photons in the element (or difference in average numbers between it and the surround) to the rms deviation from the average. For an average number $n$, the signal-to-noise ratio is $n^{1/2}$.

The threshold signal-to-noise ratio $k$ is the ratio of signal to noise required to avoid false alarms. Its value is normally about 5 and may be as low as 3 under extreme low-light conditions.

The characteristic for an ideal photon-noise-limited system is

$$nd^2C^2 = k^2$$

where $n$ is the number of photons/cm$^2$, $d$ is the diameter of test element, $C$ $(= \Delta B/B)$ is the contrast of the test element with the surround, and $k$ $(= 5)$ is the threshold signal-to-noise ratio.

The signal-to-noise ratio of a system has meaning only when the size of the test element has been specified.

The resolution of a system has meaning only when the contrast of test element has been specified.

Geometric limitations on resolution affect the high-contrast elements more than the low-contrast elements.

Bar patterns are more visible than an isolated spot whose diameter is equal to the width of the bar.

## 1.10. References

C-1. J. W. Coltman, Scintillation limitations to resolving power in imaging systems. *J. Opt. Soc. Am.* **44,** 234 (1954).

C-2. J. W. Coltman and A. E. Anderson, Noise limitations of resolving power in electronic imaging. *Proc. IRE* **48,** 858–865 (1960).

H-1. L. Hayen and R. Verbrugghe, A comparison of the signal-to-noise ratio and sensitivity of film and plumbicon camera, *J. Soc. Motion Picture Television Engrs.* **81,** 184 (1972). The authors find that the signal-to-noise ratio of a studio television picture exceeds that of 16-mm, color reversal film, ASA-125. Many of the 16-mm film clips used on television are significantly noisier than the color film used by these authors.

K-1. I. J. Kaar, The road ahead for television, *J. Soc. Motion Picture Engrs.* **32,** 18 (1939).

M-1. R. H. Morgan, Threshold visual perception and its relationship to photon fluctuation and sine-wave response, *Am. J. Roentgenol., Radium Therapy Nucl. Med.* **93**, 982–996 (1965).

R-1. A. Rose, The sensitivity of the human eye on an absolute scale, *J. Opt. Soc. Am.* **38,** 196 (1948).

S-1. R. E. Sturm and R. H. Morgan, Screen intensification systems and their limitations, *Am. J. Roentgenol. Radium Therapy* **62,** 617 (1949).

**General**

R. Clark Jones, "Quantum Efficiency of Detectors for Visible Infrared Radiation," in *Advances in Electronics and Electron Physics,* Vol. 11, pp. 87–183 (1959), Academic Press, New York.

A. Rose, "Television Camera Tubes and the Problem of Vision," in *Advances in Electronics and Electron Physics,* Vol. 1, pp. 131–166 (1948), Academic Press, New York.

A. Rose, "Quantum Effects in Human Vision," in *Advances in Biological and Medical Physics,* Vol. 5, pp. 211–242 (1957), Academic Press, New York.

O. H. Schade, The resolving-power functions and quantum processes of television cameras, *RCA Rev* **28,** 460–535 (1967).

R. E. Sturm and R. H. Morgan, Screen intensification systems and their limitations, *Am. J. Roentgenol. Radium Therapy* **62,** 617 (1949).