Statistics of Television Signals

By E. R. KRETZMER

(Manuscript received February 28, 1952)

Measurements have been made of some basic statistical quantities characterizing picture signals. These include various amplitude distributions, autocorrelation, and correlation among successive frames. The methods of measurement are described, and the results are used to estimate the amount by which the channel capacity required for television transmission may be reduced through exploitation of the statistics measured.

INTRODUCTION

One of the teachings of information theory is that most communication signals convey information at a rate well below the capacity of the channels provided for them. The excess capacity is required to accommodate the redundancy, or repeated information, which the signals contain in addition to the actual information. Removal of some of this redundancy would reduce the channel capacity required for transmission, thus opening the way for possible bandwidth reduction. In order to remove redundancy, one must first understand it; the amount and nature of the redundancy can be completely defined in terms of various statistical parameters characterizing the signal.

It has been pointed out that the existence of redundancy is particularly evident in the case of television; moreover, its elimination is highly desirable because of the large bandwith presently required for transmission. Evidence of redundancy is found in the subject matter of television—the average scene or picture. Knowing part of a picture, one can generally draw certain inferences about the remainder; or, knowing a sequence of frames, one can, on the average, make a good guess or prediction about the next frame. In either case, knowledge of the past removes uncertainty as to the future, leaving less actual information to be transmitted.

Another way of looking at this is to visualize the picture as an array of approximately 210,000 dots, 500 vertically, 420 horizontally, corresponding, respectively, to the 500 scanning lines and 420 resolvable picture elements per line of the standard television raster. Each dot can have, say, 100 distinguishable brightness values in a good-quality picture. The number of possible combinations is therefore approximately $100^{210,000}$ or $10^{420,000}$. At the usual rate of 30 frames per second it would take approximately $10^{419,991}$ years to transmit all these "pictures," which our present television system is fully prepared to transmit! The vast majority of these "pictures" will, of course, never be transmitted in this age because the average picture statistics virtually preclude the possiblity of their occurrence.

If all of the redundancy alluded to in the preceding paragraph were to be expressed in terms of statistics, the array of data would be staggering.* Redundancy encompassing even a small part of a single frame implies statistics of enormously high order because of the large number of possible past histories. The initial attention should therefore be focused on local redundancy, encompassing only a few adjoining picture elements. Accordingly, measurements have been made of the following statistical quantities.

1. *Simple probability distribution of signal amplitudes corresponding to picture brightness.* This encompasses only a single picture element, revealing the relative probabilities of this or any element's assuming the various possible brightness values, in the absence of any past-history information.

2. *Simple probability distribution of error amplitudes resulting from linear prediction of television signals.* Only the simplest type of linear prediction is considered here, so-called previous-value prediction, which predicts each picture element to have the same brightness value as the preceding one. The prediction error signal is simply the difference between the picture signal and a replica delayed by one Nyquist interval (one-half the reciprocal bandwidth or the time interval corresponding to the spacing between picture elements). The distribution of this error signal encompasses two picture elements (past history of one element) and therefore is a condensed version of the family of first-order joint probability distributions.

3. *Autocorrelation of typical pictures.* This statistical quantity is an even more streamlined version of various families of different-order joint probability distributions. Each family corresponds to just a single point on the autocorrelation curve; the ordinates of the curve represent the average correlation between picture elements spaced by various

* Complete statistics extending, say, over one frame period, would comprise one conditional probability distribution per picture element for each possible past history. With the approximate figures cited above, the number of distribution curves (many of which would be similar) is $210,000 \times 10^{419,999}$ or $10^{420,004.3}$.

distances. This correlation, say, between horizontally adjoining elements is simply the average product of the two brightness values of each pair of neighbors, relative to the average square of all brightness values.

The three quantities enumerated above contain a great deal of statistics in very compact form, but these statistics are essentially of a local and linear nature. They do not include the bulk of the large-scale redundancy, which is of a far-flung and nonlinear nature.

AUTOCORRELATION

For a function of time, f(t), the autocorrelation can be expressed as

$$\phi(\tau) = \overline{f(t)\, f(t + \tau)} \tag{1}$$

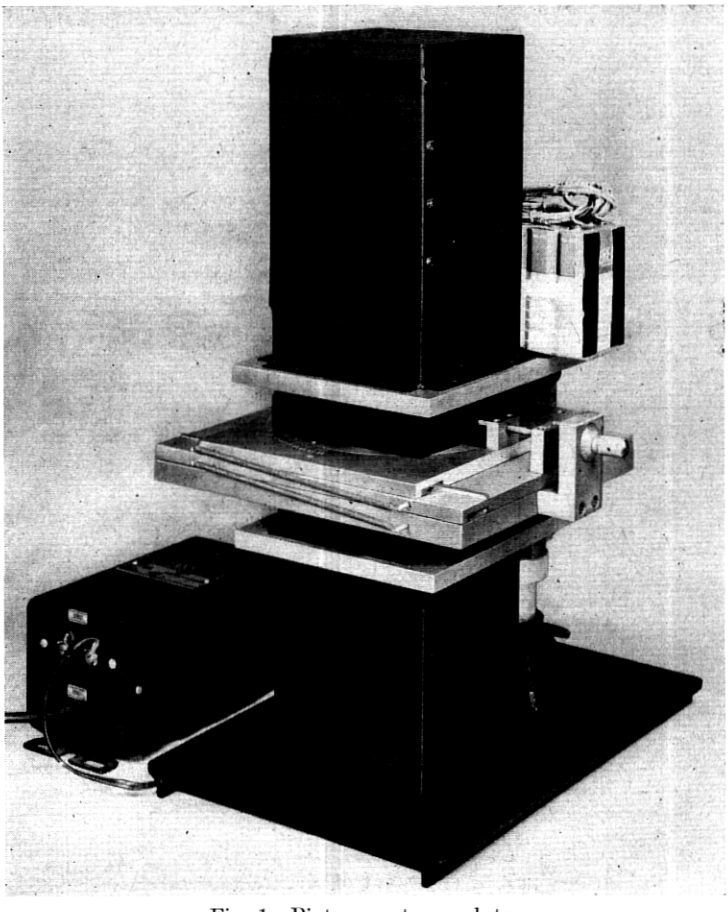

Fig. 1—Picture autocorrelator.

averaged over all time, for various values of the time shift τ. In the case of a picture transparency, the optical transmission is a function of two-dimensional space, expressible in polar coordinates as $T(s\underline{/\phi})$, and the autocorrelation can be expressed in analogous fashion. The time variable t is replaced by the space coordinate $s\underline{/\phi}$, and the correlation time shift τ is replaced by a space shift $\Delta s\underline{/\theta}$, so that the new expression is

$$\phi(\Delta s\underline{/\theta}) = \overline{T(s\underline{/\theta})\ T(s\underline{/\theta} + \Delta s\underline{/\theta})}, \tag{2}$$

averaged over as much area as practicable. This space-domain autocorrelation is much easier to measure than the time-domain autocorrelation. We need merely measure the relative optical transmission of two identical cascaded transparencies, shifted from register by a variable

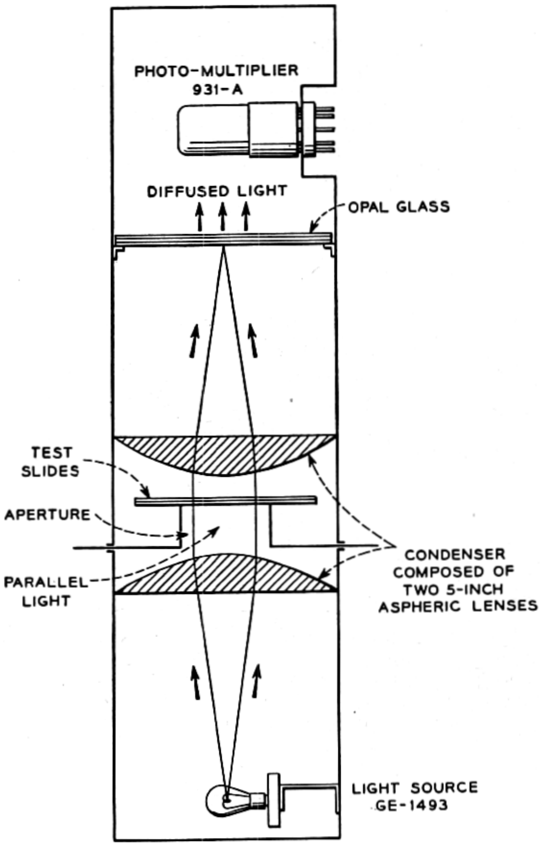

Fig. 2—Basic arrangement of picture autocorrelator.

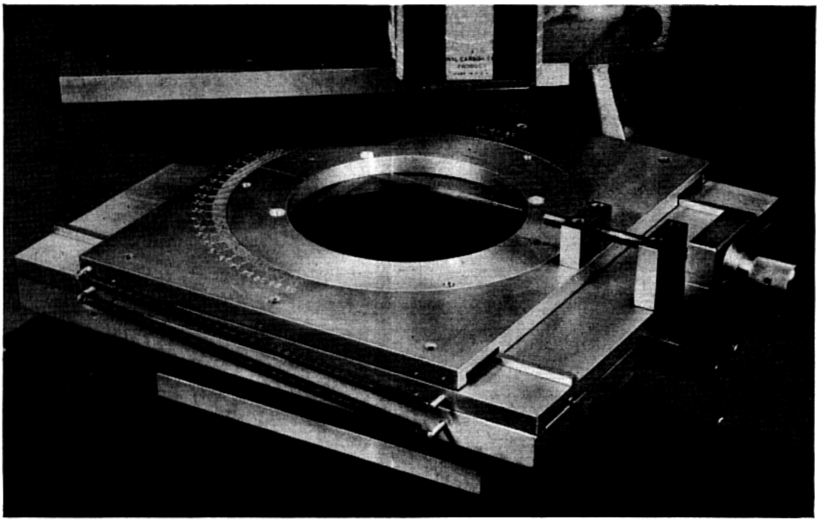

Fig. 3—Close-up view of slide holding assembly and shifting mechanism of picture autocorrelator.

amount. The averaging process is inherent in such a measurement.

The apparatus used to measure autocorrelation is shown in Figs. 1 and 2. The chamber at the bottom contains a light source of very constant intensity and a convex lens to collimate the light. The middle part, made of accurately machined aluminum, holds the two identical slides of the picture under test, and an aperture exposing a large circular area of the slides. The top chamber contains a collector lens and a photomultiplier tube which (on a microammeter not shown) gives a sensitive indication of the total light transmitted through the slides. Fig. 3 shows a close-up view of the slide-holding assembly. Two close-fitting graduated aluminum rings permit accurately determined rotation of both slides or one slide, and the micrometer drive permits translational displacements measurable to within one mil (moving the two slides by equal and opposite amounts); the separation between picture elements is approximately 7.5 mils horizontally and 5 mils vertically (for the $2\frac{1}{2}''$ by $3\frac{1}{4}''$ slide size used).

The light transmission is always a maximum when the two slides are in precise register ($\Delta s = 0$). For large shifts the transmission fluctuates about a nonzero asymptote. The nonzero asymptote results from the fact that the average transmission is always positive, and the fluctuation from the fact that large displacements introduce substantial amounts of new picture material into the aperture. Since these components tend to obscure the correlation effects, it is useful to make additional measurements which enable us to subtract them out completely. This leaves us with a 'pure' autocorrelation $A(\Delta s\underline{/\theta})$, which is then normalized so as to have a peak value of unity. It is given by

$$A(\Delta s\underline{/\theta}) = \frac{T_2\left(\pm \frac{\Delta s}{2}\underline{/\theta}\right) - T_1\left(\frac{\Delta s}{2}\underline{/\theta}\right) T_1\left(-\frac{\Delta s}{2}\underline{/\theta}\right)}{T_2(0) - T_1^2(0)}, \tag{3}$$

where $T_2\left(\frac{\pm\Delta s}{2}\underline{/\theta}\right)$ is the transmission through the two cascaded slides shifted by equal and opposite amounts $\frac{\Delta s}{2}$ at an angle θ with the horizontal, and $T_1\left(\frac{\Delta s}{2}\underline{/\theta}\right)$ is the transmission of a single slide with displacement $\frac{\Delta s}{2}$ at the same angle θ.

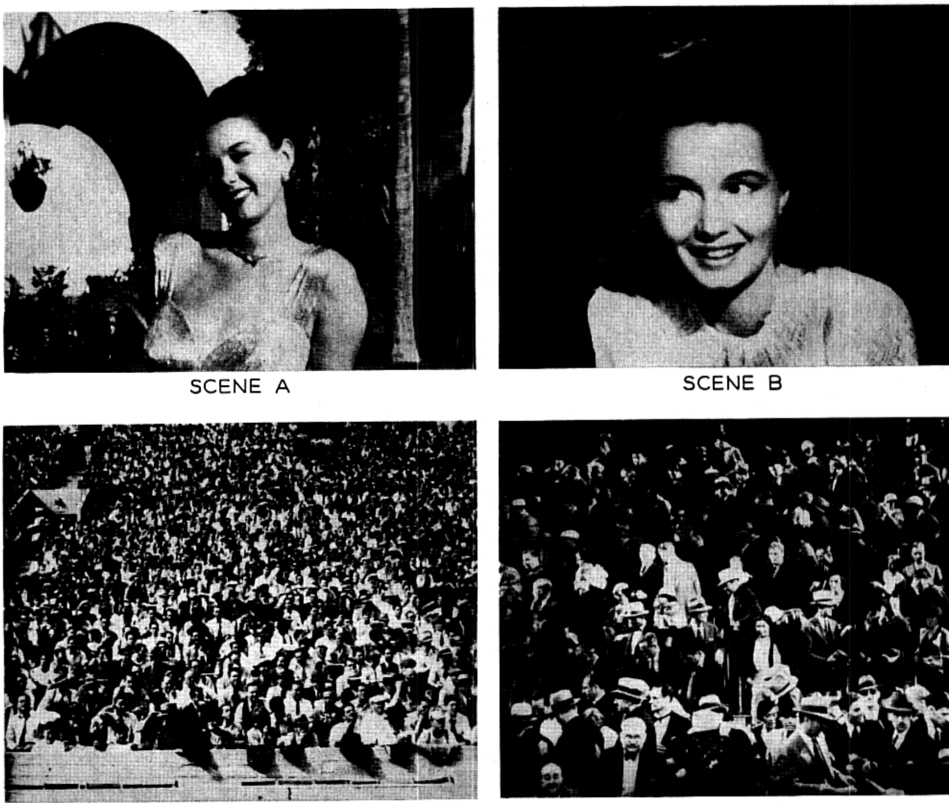

Fig. 4—Test pictures whose statistics are included in this article.

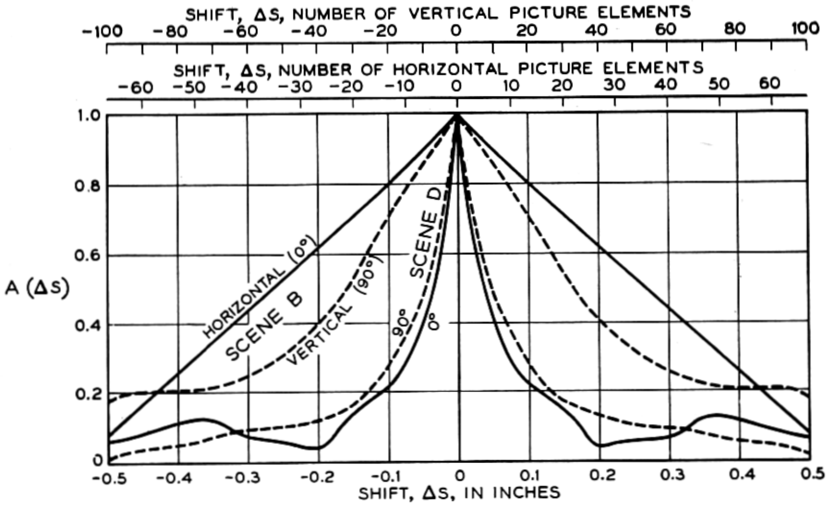

Fig. 5—Plots of autocorrelation in horizontal and vertical directions for two pictures.

Fig. 4 shows some pictures for which autocorrelation measurements have been made. The results can be presented in the various ways shown in Figs. 5, 6, and 7. Fig. 5 shows conventional plots of A versus Δs in the horizontal and vertical directions. Scene B is seen to have more correlation than Scene D, and curve shapes range from remarkably linear to somewhat like exponential. Fig. 6, giving contours of constant autocorrelation, brings out the variation with the angle θ. Scene A happens to have its greatest correlation in the vertical direction, but that was not found to be a general rule by any means; Scene B, for example, has its greatest correlation in the horizontal direction. No preferred directions appear to exist in general. In Fig. 7 attention is focused on the more local correlation, for small values of Δs. The average correlation among horizontally adjoining picture elements, designated by A_{10}, is seen to be approximately 0.99 for Scene B and only 0.75 for Scene C. A_{20} denotes the correlation for a horizontal spacing of two picture elements while A_{01} denotes the correlation among vertically adjoining picture elements.

It should be pointed out that the pictures which gave the above results were not band-limited to the standard 4-mc resolution. However, before the results were used quantitatively, the proper band limitation was applied mathematically. This has the effect of rounding off the peaks of the curves, decreasing the autocorrelation drop within the first Nyquist interval by up to approximately 24 per cent.

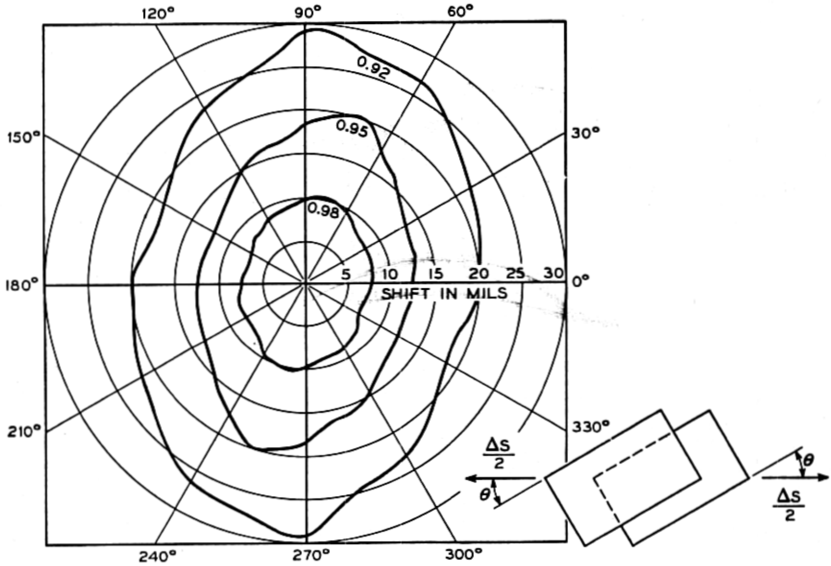

Fig. 6—Contours of constant autocorrelation for Scene A. In general there are no preferred directions of correlation.

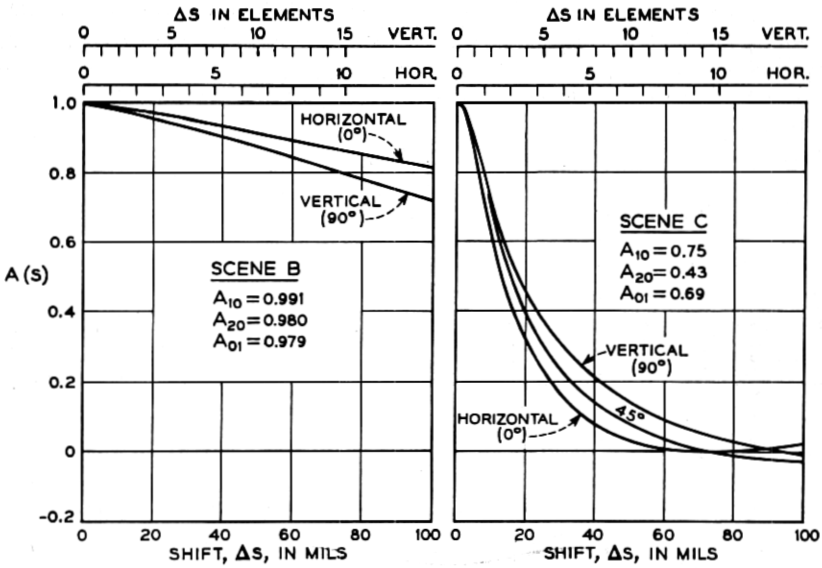

Fig. 7—Plots of autocorrelation for small shifts. A_{10} is the autocorrelation for a shift of one horizontal elemental distance, A_{20} for two horizontal elemental distances, and A_{01} for one vertical elemental distance. Alternatively A_{10} may be described as the average correlation between horizontally adjoining elements, etc.

PROBABILITY DISTRIBUTIONS

A probability distribution of amplitudes is generally shown as a plot of probability density versus signal amplitude. Probability density, say, corresponding to amplitude x_1, is the probability of finding the signal amplitude between x_1 and $x_1 + dx$, divided by the differential amplitude increment dx. Conversely, the probability of finding the signal amplitude between x_1 and $x_1 + dx$ is given by $p(x_1)dx$, $p(x)$ being the probability density corresponding to amplitude x.

If a cathode-ray spot is deflected, say horizontally, by the signal in question, its average dwell time at any point is directly proportional to the corresponding probability density. In the optical system shown in Fig. 8, a cylindrical lens maps each point into a vertical line which is

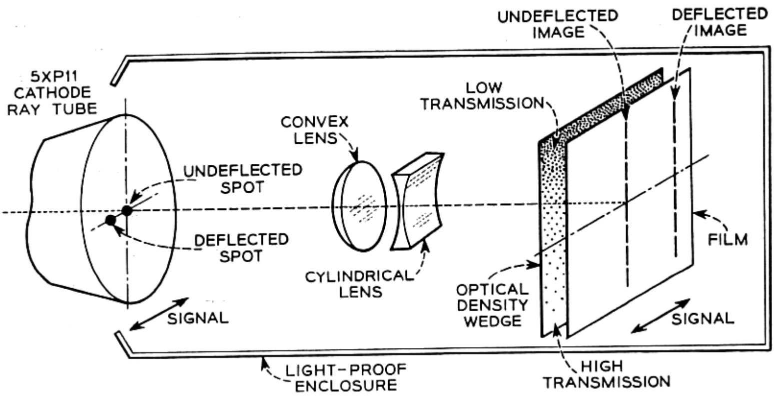

Fig. 8—Basic arrangement of probabiloscope.

then tapered in intensity by an optical density wedge before reaching a high-contrast photographic film. Depending on the dwell time at any amplitude level, the corresponding tapered line has enough average intensity to blacken the film up to a certain level. This level is proportional to log $p(x)$, since the density wedge is tapered exponentially so that the intensity of each tapered line of light reaching the film diminishes, say, by a factor of ten for each inch we travel up the line. The film in effect traces out a contour of constant exposure.

Two or three iterated photographic printings increase the effective gamma sufficiently to yield a contour of ample sharpness. This contour is then changed to a sharp line by a simple dark room trick: while the film is in the development tray, already fully developed, it is momentarily exposed to light. The blackened portion of the film is unaffected, the clear portion is fully blackened, while the transition contour, being partly opaque, is not fully blackened. By printing from this film we then obtain a well-defined black-on-white curve of $p(x)$ versus x on a logarithmic probability scale. The logarithmic scale has the advantage of making the curve shape independent of exposure length and giving uniform relative accuracy over the entire range.

Fig. 9 shows some typical results obtained by means of the "probabiloscope." The two small curves are distributions of two different still pictures. The left-hand end corresponds to black, the right-hand end to peak white; the blanking intervals (slightly blacker than black) cause the peaks at the extreme left. (The signals did not contain any synchronizing pulses.) The tall and slender curve at the right of Fig. 9 is the distribution of errors resulting from previous-value prediction of one of the pictures in Fig. 4. The peak corresponds to zero error which is seen to be most probable, as it should be if the prediction criterion is good. Increasingly larger errors are increasingly improbable or rare. The six decades of probability density spanned by the curve were obtained in three separate exposures and subsequently joined, since stray

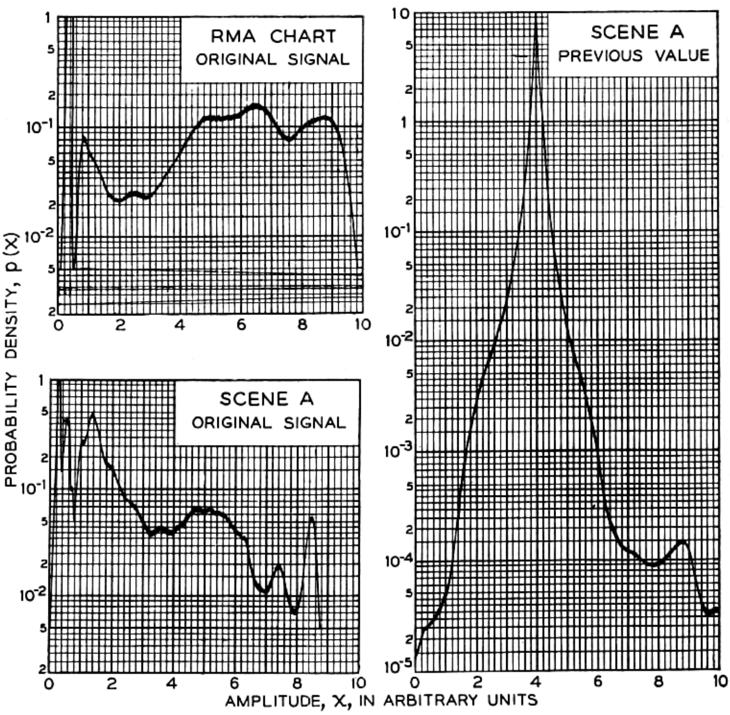

Fig. 9—Typical probability distributions as obtained from the probabiloscope. Curves at left are for video signals; right-hand curve is for difference between video signal and delayed replica.

light limits the useful range of the probabiloscope to approximately two decades. In obtaining those sections of the curve corresponding to the few and far-between large errors, a long exposure was used and the cathode-ray beam was blanked whenever passing through the range of zero or small errors. The vertical scale on all curves is determined solely by the density taper of the optical density wedge. If this scale is to represent true probability density, instead of a proportional quantity, it should be shifted up or down so as to make the area under the curve equal to unity.

APPLICATION OF RESULTS

The statistics measured can be put to various uses, such as in the design of better predicting or coding schemes. The most interesting application is probably in estimating the reduction in channel capacity which the measured statistics show to be theoretically possible. In other words, the results can give us various lower bounds to the redundancy of television signals.

For the sake of illustration, suppose that the signal is quantized into 64 amplitude levels. An ordinary television channel assumes all 64 levels to be equally likely, hence is prepared to accommodate $\log_2 64$ or 6 bits per sample. But the simple amplitude distribution of the signal is not flat, so that all 64 levels are *not* equally likely. The maximum possible associated average information content per sample is given by

$$H_{\max} = \sum_{i}^{64} p_i \log p_i \,, \tag{4}$$

where p_i is the simple probability of the signal's falling into the ith level. Since the 64 p_i's are unequal, $H_{\max}$ is necessarily less than 6 bits. For all available data the average value of $H_{\max}$ turns out to be approximately 5 bits, indicating a one-bit redundancy. The latter figure is essentially independent of quantization.

The prediction error signal still contains all the useful picture information. The maximum possible information content per sample (maximum in that all samples are assumed to be completely independent) is still given by (4) but in this case the 64 values of p_i are obtained from the peaked error distribution. The average* result from all available data turns out to be approximately 3.4 bits below the 6-bit ceiling, show-

* This average was computed by averaging the various redundancy values obtained for the individual pictures, rather than averaging all statistical data and then finding one corresponding average redundancy. The average computed here is more favorable and can be realized only if optimum coding is performed on a short-term basis rather than on the basis of one set of long-term statistics.

ing that the original signal must have contained at least 3.4 bits of redundancy.

The autocorrelation can also furnish a lower bound to the redundancy, as has been pointed out by P. Elias in his Letter to the Editor of the *Proceedings of the I.R.E.* for July, 1951. If, for example, the correlation A_{10}, between horizontally adjoining picture elements, is high, the corresponding lower-bound redundancy is very roughly equal to

$$R \approx -\tfrac{1}{2} \log_2 (1 - A_{10}) \text{ bits/sample.} \tag{5}$$

Alternatively, taking the Fourier transform of the autocorrelation yields the power spectrum $P(f)$, from which we can find the lower-bound redundancy through the relation

$$R = \frac{1}{2W} \int_0^W \log_2 P(f)\, df + \tfrac{1}{2} \log_2 W + \log_2 K \text{ bits/sample,} \tag{6}$$

where W = bandwidth in cps, and $\frac{1}{K} = \int_0^W P(f)\, df.$

Using either method, one obtains approximately 2.4 bits for the average* of the available data. This is an approximate bound, in that it applies strictly only to functions having gaussian amplitude distributions.

Suppose, then, that we have exposed an average redundancy of at least 3 bits per sample. This means a potential 3-bit reduction in the channel capacity required for television transmission. In a 6-bit system (64 amplitude levels) this means a 50 per cent reduction, and hence a potential halving of the bandwidth with the aid of an ideal coding scheme. It is true that the decorrelated signal is somewhat "frail," i.e., vulnerable to interference, so that it might be desirable to use a "rugged" system of the PCM variety for transmission. Thus, if a Shannon-Fano code were used, the 3-bit decorrelation should enable us to send television by an average of 3 on-off pulses per picture sample rather than 6. This represents a two-to-one saving over the usual PCM bandwidth. More spectacular reductions are likely to be achievable only by tapping the large-scale redundancies mentioned earlier.

FRAME-TO-FRAME CORRELATION

There is, of course, a great deal of interest in the possibility of utilizing the similarity between successive frames. Accordingly, adjacent-frame

* See previous footnote.

correlation was measured for two typical motion-picture films, by means of the apparatus described in the section on autocorrelation.* The results were 0.80 and 0.86, after correction for the 4-mc bandwidth limitation. This means that "previous-frame" prediction can remove only slightly more than one bit of redundancy per sample. More complicated schemes would presumably be more successful in taking advantage of the large frame-to-frame redundancy which undoubtedly exists.

ACKNOWLEDGMENT

Many of the ideas expressed in this paper are due to B. M. Oliver, whose resourcefulness is hereby gratefully acknowledged.

* The expression used in evaluating the correlation between frame 1 and frame 2 (any two frames) is

$$C_{12} = \frac{T_{12} - T_1^2}{T_{11} - T_1^2}, \tag{7}$$

where T_{12} is the optical transmission of frames 1 and 2 in cascade, T_1 is the average of the individual transmission of frames 1 and 2, and T_{11} is the average of the transmissions of two cascaded slides of frame 1 and two cascaded slides of frame 2, respectively. In all cascade transmission measurements, the two frames must be in precise register.