621.397 : 621.391

Paper No. 1401
RADIO SECTION

SOME POSSIBILITIES FOR THE COMPRESSION OF TELEVISION SIGNALS BY RECODING

By E. C. CHERRY, D.Sc., and G. G. GOURIET, Associate Members.

(*The paper was first received* 17*th January, and in revised form* 9*th July*, 1952.)

SUMMARY

The paper describes first the excessive redundancy in television signals, using the term in the sense of communication theory. The information in a picture is largely contained in the edges and boundaries, but in practice these occupy a very small part of the total area. The unequal probabilities of signal elements lying in boundary and smooth surface regions are considered, together with the transition probabilities between adjacent elements. Compression is shown to be possible by making such probabilities more nearly equal, using a system of recoding.

Such recoding may be achieved by a variable-velocity scanning system if the scan velocity is made high for all regions of the picture having low picture-detail and low for regions having high picture-detail. Picture-detail is defined precisely and measured values are given from observations on working programme material.

Finally, compression is shown to be theoretically attainable in both the vertical and the horizontal directions. Correlation between successive line-scan waveforms is interpreted in terms of transition probabilities, and again a redundancy is seen to exist. Suggestions are made for reducing this by storage of every successive frame and scanning alternately in the horizontal and vertical directions, in both cases using a variable-velocity system. Finally, the paper should not be read as describing an existing working system, but purely as a discussion of possibilities.

(1) INTRODUCTION

The principles of bandwidth compression presented in the paper are applicable to many systems of communication, in particular to telephony and television; the latter is the more topical subject and attention will here be confined to this, although the compression of telephony channels may soon become a matter of practical urgency. The approach which has been adopted is that of statistical communication theory, using the work of Shannon;[1] the potential compressibility of television signals is shown to depend upon their excessive redundancy. Most simply, the redundancy arises in a signal, such as a television waveform, by virtue of the fact that if certain parts of the signal are known other parts may be "guessed" with more or less accuracy. The greater this predictability, the more the redundancy. Reduction of this redundancy may be achieved by recoding the signal so as to change its statistical structure. Such recoding, in effect, achieves bandwidth compression by taking more time over those elements of the signal which contain the most information and reducing the time for those which contain less.

In order to comprehend the value of recoding and the mechanism of compression it is essential to regard signals in terms of probabilities. Shannon's own illustration is as simple as any for emphasizing this point of view. Consider a message written, say, in English. There are two kinds of probability which are relevant; first, the various letters of the alphabet have different probabilities of occurring, those of highest probability containing least information and vice versa. Morse appreciated this when designing his code,[2] and he attached the shortest symbols to the most frequently occurring letters, thereby achieving coded messages which contain a minimum number of symbols (dots and dashes). Such a recoding also renders the probabilities of the symbols more nearly equal; ideally it would be best to make them quite equal. However, there is another most important factor in our written message, namely that the successive letters are not independent.[3] Thus, knowing one letter in a word we have a certain chance of guessing the next; more precisely, the next letter is represented by a transition probability distribution over the letters of the alphabet. This constitutes a di-gram, or 2-letter, transition probability, $p_i(j)$. Furthermore, there are tri-gram . . . *n*-gram transition probabilities. Again, it is possible to smooth these transition probabilities by recoding.

It has been shown[1] that such redundancy is essential for combating noise, but that waste of channel capacity (e.g. bandwidth) results if the redundancy exceeds some value which depends upon the noise level. Considerable bandwidth may be saved by reducing the redundancy in systems where it exists in excess, particularly in telephony and television, and to achieve this, coding is required:

(*a*) To equalize transition probabilities between the successive signal elements or "symbols." If these are completely equalized the symbols become independent of one another and equally probable. In general, this is equivalent to:

(*b*) Equalizing the probabilities of occurrence of all symbols admitted in the transmission. On the other hand, (*a*) does not follow from (*b*).

In the paper a system is described in principle which changes the statistical structure of television signals by means of velocity-modulation of the scanning beam. Such velocity-modulation is quite distinct from that described by Bedford and Puckle,[4] since the scan velocity is made to depend, not upon the instantaneous picture brightness, but upon its detail from point to point (where this quality has yet to be defined). In this way a longer time is occupied in scanning the areas of considerable picture detail and a much shorter time in those areas of zero or little detail.

(2) THE CHARACTER OF TELEVISION SIGNALS

(2.1) The Transmission of Simple Patterns

The traditional view of television definition and the assessment of the information-transmitting qualities of a channel are based upon the chequerboard of alternate black and white picture-points. With the introduction of communication theory, for the purpose of bandwidth compression, this view has to be modified. Let us consider first how to describe the simple pattern shown in Fig. 1(*a*), for example over a telephone, to a friend who is attempting to reconstruct it from our description. We could, of course, imagine the pattern to be covered by a chequerboard mesh and describe the pattern picture-point by picture-point, starting at the top left-hand corner and proceeding scan-wise to the bottom. But it would clearly be far simpler to take the first

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Dr. Cherry is in the Department of Electrical Engineering, Imperial College of Science and Technology.
Mr. Gouriet is with the British Broadcasting Corporation.

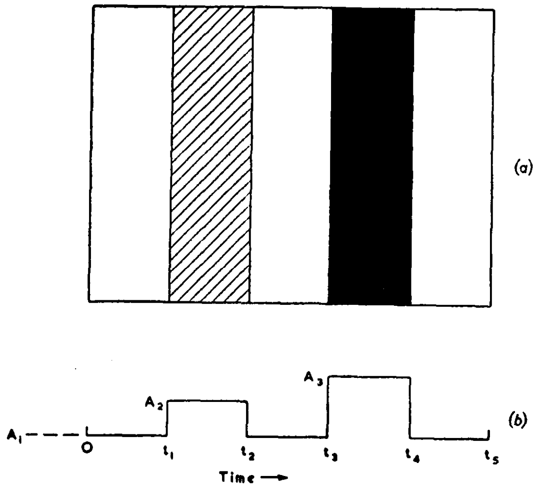

Fig. 1.—Description of simple pattern.

(a) Simple test pattern.
(b) Line-scan waveform.

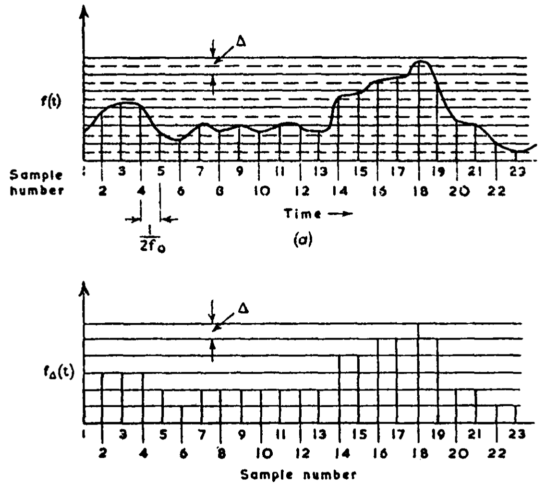

Fig. 2.—Quantized television signal.

(a) Sampled television waveform.
(b) Quantized pulse form.

scan-line and merely to describe the co-ordinates of the edges of the pattern and their amplitudes [Fig. 1(b)], and then to say "continue with every other line in the same way." A simple set of number pairs would suffice, as in this example:

$$\text{Line (1):}\quad S_1 = (A_1, t_1)(A_2, t_2)(A_1, t_3)(A_3, t_4), (A_1, t_5) \quad . \quad . \quad (1)$$

However, it might be argued that real television pictures are more complex and that the chequerboard method might then not be so absurd. But it happens that pictures having the complexity of the chequerboard never occur, i.e. they have zero probability. Measurements have shown that the complexity of practical television pictures is of the order of 3% and only on rare occasions exceeds 8% of that of the chequerboard. The term "complexity" is defined more precisely in the succeeding Section, and details of the measurements are given in Section 2.3.

If a system of bandwidth compression is required, it would appear that a certain risk must be taken that test patterns could be devised which would be incapable of transmission; but with working television programmes scenes having similar characteristics never, in fact, occur.

(2.2) The Information in a Television Picture

The simple example of Fig. 1 serves to illustrate one important fact, namely that the boundaries or edges contained in a picture are highly significant. This is apparent from common experience; for example, pen-and-ink drawings may be as recognizable as photographs. Again, if a television picture-signal is differentiated before being presented on the viewing-tube screen it still provides a recognizable picture. We may say, somewhat loosely, that the essential information in a picture is contained in the edges and boundaries. But it will be necessary to relate this use of the word "information" to its technical use in communication theory before considering the problem of recoding it.

Everyday observation of scenes, faces or landscapes illustrates the fact that the total visual area occupied by edges and boundaries is relatively small, but by far the larger proportion of the area is occupied by uniform or near-uniform tones and brightness levels, i.e. by steady-state signals. Thus, the probability of edges is far less than the probability of steady-state levels.

Consider a portion of one line-scan of some television scene, as shown in Fig. 2(a). It might be argued that absolute steady-state level is, in fact, a rare occurrence. However, recent work has shown[5] that amplitude quantization of television signals is a real practical possibility, and that remarkably clear pictures can be obtained with relatively few quantum levels. The line-scan waveform $f(t)$ in Fig. 2(a) is shown sampled at the usual regular intervals of $1/2f_0$, where f_0 is the limiting upper frequency of the channel bandwidth. These samples represent the amplitudes of the pulses required to convey the waveform by pulse amplitude modulation, and if this pulse signal is passed through an ideal filter, with cut-off frequency f_0, the output signal will correspond to the original waveform $f(t)$. Again, Fig. 2(a) is shown marked (for example) into a series of seven equal-amplitude (quantized) levels, including zero, and a quantized pulse signal $f_\Delta(t)$ results if the successive samples are allowed to adopt their nearest quantized amplitudes, as shown in Fig. 2(b).

In this Figure the successive samples are labelled 1, 2, 3 . . . n, etc. Those samples of the original waveform $f(t)$ which lay between two adjacent quantum levels but showed a certain degree of ripple or amplitude variation (such as samples 7–13) now become quite equal after quantization. Thus the samples, which represent the successive "symbols" of the line-scan waveform, fall into two classes:

(a) Those forming steady-state regions, where two or more successive samples are identical in amplitude (as 7–13) in Fig. 2(b).

(b) Those whose quantum level differs from the preceding sample level in the direction of scanning. These constitute the picture edges (as in 1–2, 4–5 or 13–14) in Fig. 2(b).

It is these two classes of sample which have very differing probabilities; as mentioned in Section 2.1, the edge samples occur only seldom, whilst the steady-state samples preponderate.

Given any one sample of known amplitude (quantum level i) the amplitude of the next one (quantum level j) is given by a transition probability $p_i(j)$. For half-tone pictures, as opposed to full black-and-white drawings, it may first be assumed that $p_i(j)$ is highest for equal quantum levels, that it is somewhat less for a quantum jump of one level, $\pm\,\Delta$, less still for a double jump, $\pm\,2\Delta$, and so on. The actual values of these probabilities can be determined only from an ensemble of such waveforms from the same source; in practice, this means that the probabilities are obtained from statistical measurements on the information source (i.e. the television transmitter producing the pictures). Shannon has derived a formula[1] for the average information, H, from a source possessing a known transition probability structure:

$$H \propto -\sum_{i,} p_i p_i(j) \log p_i(j) \quad . \quad . \quad . \quad . \quad (2)$$

where p_i is the probability of any symbol i and $p_i(j)$ the transition probability of symbol j following. On the assumption that, on an average, all quantum levels are equally likely, p_i will be a constant. The information conveyed by any one sample-pair transition is $\log p_i(j)$, whilst eqn. (2) is the average of this over a very long series of samples. Clearly, on the basis of the assumptions we have made the information conveyed by the maximum quantum jump (full black to full white) is greatest, whereas the information conveyed by zero jump (steady state) is least. The problem of signal compression thus reduces to that of recoding the signal so that in a given interval of time there are produced far more larger quantum jumps, in relation to the number of smaller or zero jumps, than in the original signal, i.e. the transition probabilities $p_i(j)$ must be made more nearly equal.

We are now in a position to define a simple measurable quantity which may be called the "picture detail" which, referred to any specific region of the picture, is some measure of the information conveyed. This picture detail is the sum of the number of samples, 1, 2, 3 . . . n contained in the region (say part of a line-scan) weighted according to the magnitudes of their quantum jumps, Δ_1, Δ_2 . . . Δ_n, and given by

$$D_n = \sum_n \Delta n \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where, it must be remembered, we have at present assumed the probability $p_i(j)$ of any such quantum jump to decrease linearly with the magnitude Δn.

In the limit, as the quantum-level differences approach zero, this becomes a measure of the picture detail of the continuous waveform. If e is the amplitude of the waveform as a function of the scanning distance, x, then over any scan length, l, the picture detail may be defined as

$$D_l = \int_l \left|\frac{de}{dx}\right| dx \quad . \quad . \quad . \quad . \quad . \quad (4)$$

which, for a fixed length l, is the absolute sum of all the amplitude variations along the length l. Fig. 3 illustrates this. Over

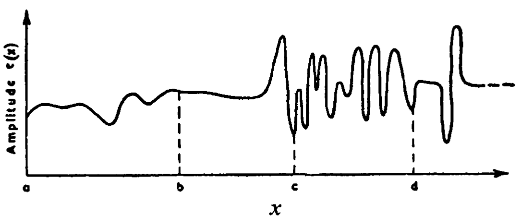

Fig. 3.—Picture detail, $\int \left|\frac{dF}{da}\right| dx$.

a–b .. Region of low detail.
c–d .. Region of high detail.

regions such as a–b the picture detail is low, since both the amplitude and the frequency are low; over regions such as c–d it is higher. Clearly, a steady-state d.c. level contains zero picture-detail, whereas an oscillation of maximum frequency f_0 at full modulation depth A_{max} contains maximum picture-detail and maximum information. A waveform of this latter type would result from scanning a black-and-white chequerboard test-pattern; thus, such a test-pattern represents 100% picture detail, a degree of detail which a practical television system is never called upon to produce.

It must be emphasized that neither eqns. (3) nor (4) are, in general, a measure of information in Shannon's sense. Thus, the expression "picture detail" is preferred to "information." However, there exists one particular distribution $p_i(j)$ of pulse amplitudes for which the picture-detail function D_l becomes identical with Shannon's information formula (2). (See Appendix 9.)

(2.3) Measurement of Picture Detail

Apparatus suitable for measuring picture detail as defined by eqn. (4) has recently been described,[6] together with the results of measurements. For completeness a brief description of the method and results will be included here.

Picture detail, as defined by eqn. (4), is considered as a function of an arbitrary length of scan. For a television picture the significant length of scan is obviously that corresponding to one complete picture. Furthermore, for measurement purposes we require to know, not the absolute values of eqn. (4), but the ratio of the value for a given picture to the value corresponding to a picture-element chequerboard of equivalent peak amplitude. It is therefore convenient to define a measurable quantity, the "picture-detail factor," D_F, in terms of the average value of $|de/dx|$ over one complete picture scan of length L, thus:

$$D_F = \frac{1}{L}\int_L \left|\frac{de}{dx}\right| dx \quad . \quad . \quad . \quad . \quad . \quad (5)$$

If the unit of length is chosen as one picture element and the maximum value of e is taken as unity, the value of D_F corresponding with a picture-element chequerboard of maximum amplitude will be unity.

It may readily be seen that for a constant velocity of scan this may also be written:

$$D_F = \frac{1}{T_0}\int_{T_0} \left|\frac{de}{dt}\right| dt \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where T_0 is the duration of one scan, and e is the voltage output of the camera.

The quantity D_F as here defined has been measured for a wide variety of scenes, using a 405-line picture channel with a video bandwidth, f_0, of 3·0 Mc/s. The electrical analogue of eqn. (6) is very straightforward and involves only a simple differentiating circuit followed by a full-wave linear rectifier, thus giving $|de/dt|$. No integration is necessary, other than that provided by a d.c. meter reading the rectified output signal.

The measuring apparatus was calibrated for 100% picture detail by injecting a 3-Mc/s sine wave with an amplitude corresponding with the maximum amplitude of picture signal. With the output of the picture channel connected in place of the calibrating signal, picture detail was then measured directly by noting the meter reading as a fraction of the maximum reading.

During the measurements the error due to the presence of synchronizing signals and noise was eliminated by setting the meter to read zero for a blank scene containing zero information. Various tests were carried out to ensure that the apparatus was performing accurately, one of which was to mask off areas of a regular pattern used as a test object and to confirm that the reading dropped by an amount proportional to the masked area.

The measurements were carried out using a film scanner and a modern type of studio camera chosen for good resolving power. The former equipment was known to produce the full depth of modulation up to the 3-Mc/s limit of the video band, whilst the camera produced 50% modulation when the finest detail was situated in the centre of the field. This performance was considered to be typical of the best cameras in present-day use.

Photographic transparencies of high quality were used as test subjects for the camera, and 35-mm film for the flying-spot scanner.

The results of the measurements showed that the average picture-detail factor obtained with the camera, for a variety of typical scenes, was 1·8% of the maximum previously defined. The greatest value obtained with normal scenes, as opposed to test cards, was 2·4%. In this instance the scene was an elaborately furnished interior, the photographic reproduction of which contained a considerable amount of fine detail with a high contrast ratio. In all cases the quality of the television picture was equal to the highest present-day standards.

Observations carried out using the flying-spot 35-mm film scanner revealed an average figure of slightly less than 3% and a maximum of 8%, the latter occurring on one occasion only—during a newsreel. The scene which gave rise to this peak measurement was a packed grand-stand at a sports stadium, and the entire picture comprised little more than white dots representing the distant faces of spectators. Only on rare occasions did the reading exceed 5%.

Further measurements over a long period of time will be necessary to establish the frequency of occurrence of various levels of picture detail in normal programme material, but the results quoted indicate that, in practice, picture detail is seldom likely to exceed about 5% of the theoretical maximum, and a figure of 3% is consistent with a very satisfactory standard of broadcasting.

(3) RECODING BY FEEDBACK

(3.1) Velocity Modulation as a Function of Picture Detail

It is now proposed to examine in detail the result of causing the velocity of the camera scanning beam to vary inversely with what might be termed the instantaneous picture-detail, i.e.

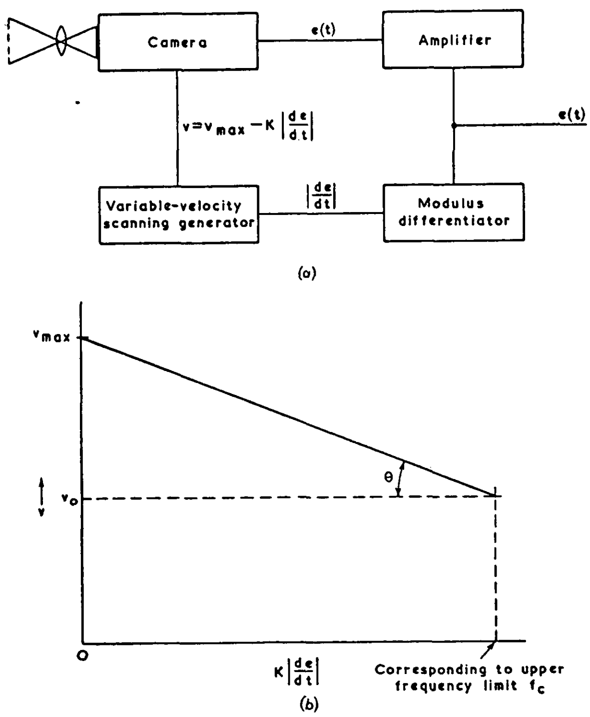

Fig. 4.—Variable-velocity system.

(a) Block schematic of feedback.
(b) Scanning characteristic.

$|de/dt|$. Fig. 4(a) shows a block schematic; with this arrangement the scanning velocity is made equal to

$$v = v_{max} - K\left|\frac{de}{dt}\right| \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where v_{max} is the maximum velocity of scan and K is a constant of appropriate dimensions. The resulting characteristic is illustrated in Fig. 4(b).

It will be noted that, unlike the velocity-scanning system proposed by Bedford and Puckle,[4] this arrangement constitutes a closed feedback loop; i.e. the beam velocity is determined by a parameter which is itself a function of the velocity.

We will assume that light from the scene falls on to the photosensitive surface of the camera and there gives rise to a charge pattern, the density of which at any given point corresponds to the light intensity at that point. Fig. 5(a) illustrates a portion of

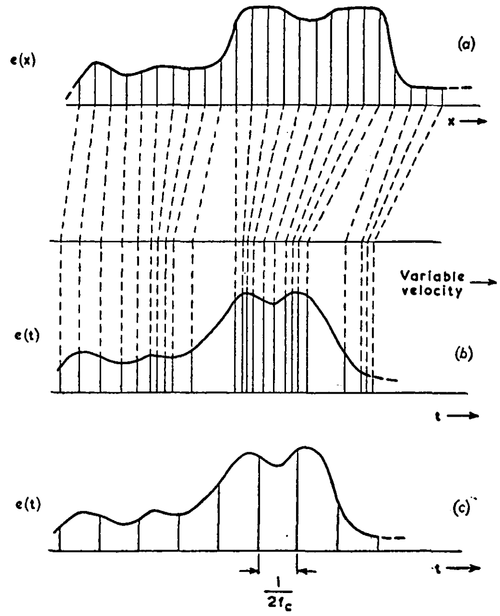

Fig. 5.—A television waveform.

(a) Stored on the camera mosaic.
(b) As a signal after velocity-modulated scanning.
(c) Sampled for transmission in a bandwidth f_c.

such a line-charge pattern, shown sampled at the usual picture-point intervals.

Let the gradient of the charge pattern which corresponds to the light image be de/dx, where x is the distance along a horizontal scanning line. For the scanning velocity we have $v = dx/dt$, and we may therefore write

$$\frac{de}{dt} = v\frac{de}{dx} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

and

$$\left|\frac{de}{dt}\right| = v\left|\frac{de}{dx}\right| \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Now let $|de/dt|$ be used to control v in a linear manner according to the characteristic shown in Fig. 4(b), so that

$$v = v_{max} - K\left|\frac{de}{dt}\right| \quad . \quad . \quad . \quad . \quad . \quad (10)$$

in which K has the dimensions of length per unit voltage. The output signal will now be as shown in Fig. 5(b), with the originally uniformly spaced samples modulated in position.

From eqns. (7) and (8) we have

$$v = v_{max} - Kv\left|\frac{de}{dx}\right| \quad . \quad . \quad . \quad . \quad . \quad (11)$$

whence

$$v = v_{max}\frac{1}{1 + K|de/dx|} \quad . \quad . \quad . \quad . \quad (12)$$

By writing

$$\left|\frac{de}{dt}\right| = v\frac{de}{|dx|} = v_{max}\frac{|de/dx|}{1 + K|de/dx|} \quad . \quad . \quad . \quad (13)$$

we see that if the product $K|de/dx|$ is large compared with unity, the modulus of the slope of the output waveform will approximate to a constant, v_{max}/K. Such a waveform is shown in Fig. 6, and its characteristics will be referred to later.

As discussed in Section 1, the object of the experiment is to

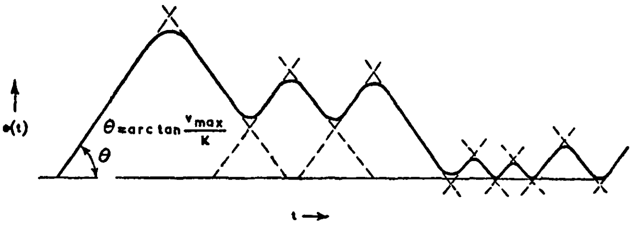

Fig. 6.—Waveform for which the modulus slope tends to a constant.

reduce the bandwidth-time product required for the transmission of a television image by arranging to scan quickly those areas which contain little detail. For a given bandwidth the time taken to scan a complete picture will be reduced, and the reduction in time is, of course, an exact measure of the possible saving in bandwidth. In order to deduce the saving we may write eqn. (12) as

$$\frac{v_{max}}{v} = 1 + K\left|\frac{de}{dx}\right| \quad . \quad . \quad . \quad . \quad . \quad (14)$$

The factor v_{max}/v will vary with the position of the scanning beam, and the average value over one complete scan of length L will be

$$\left(\frac{v_{max}}{v}\right)_{av} = 1 + \frac{K}{L}\int_L \left|\frac{de}{dx}\right| dx \quad . \quad . \quad . \quad . \quad (15)$$

In Section 2.3

$$D_F = \frac{1}{L}\int_L \left|\frac{de}{dx}\right| dx$$

was defined as the picture-detail factor, for which normalized measurements have been quoted. Again, assuming the maximum value of $|de/dx|$ to be unity, we may write eqn. (15) as

$$\left(\frac{v_{max}}{v}\right)_{av} = 1 + KD_F \quad . \quad . \quad . \quad . \quad (16)$$

in which D_F is the picture-detail factor as measured.

In order to determine the values of v_{max} and K we shall again refer to eqn. (12). The minimum velocity of scan, v_0, will result when $|de/dx|$ is a maximum, i.e. unity according to our previous definition. It will occur when the finest detail is presented at full contrast, e.g. a portion of a chequerboard pattern with alternate black and white picture elements. It should be noted here that similar fine detail of lower contrast ratio (e.g. see Fig. 7) has a proportionately smaller value of $|de/dx|$ and therefore would result in a greater velocity of scan. However, if the eventual compressed bandwidth is not to exceed an upper limit, f_c, this scanning velocity must not exceed $2f_c$ elements/sec over such regions of fine detail, and for the present we shall therefore assume that all changes are between black and white, as in a pen-and-ink drawing. The general case of half-tone reproduction will be dealt with in Section 3.2.

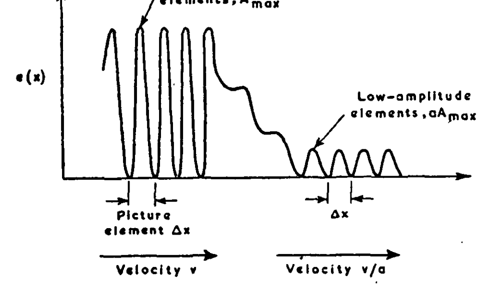

Fig. 7.—Velocity variation with change of amplitude.

Let v_{max} be n times the minimum velocity v_0 and let $|de/dx|$ equal unity, so that eqn. (12) becomes

$$\frac{nv_0}{1 + \bar{K}} = v_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

where $\bar{K}$ is the numerical value of K which results when the latter is multiplied by unit voltage per element,

$$\left.\begin{array}{ll} \text{whence} & \bar{K} = (n - 1) \\ \text{or} & K = (n - 1) \text{ elements/volt} \end{array}\right\} \quad . \quad . \quad . \quad (18)$$

Substituting these values for K and v_0 in eqn. (11) we obtain

$$\left(\frac{nv_0}{v}\right)_{av} = 1 + (n - 1)D_F$$

$$\text{or} \qquad \left(\frac{v_0}{v}\right)_{av} = \frac{1 + (n - 1)D_F}{n} \quad . \quad . \quad . \quad . \quad (19)$$

Since the time taken to complete a scan is inversely proportional to the velocity averaged over the scan, we may write eqn. (19) as

$$\frac{T}{T_0} = \frac{1 + (n - 1)D_F}{n} \quad . \quad . \quad . \quad . \quad (20)$$

in which T and T_0 are the times taken for one complete scan with and without velocity modulation respectively, and for a similar bandwidth, f_0.

This, then, is the factor by which bandwidth could be saved in the case of a pen-and-ink picture comprising two discrete amplitude levels. Thus, if the reduced time, T, for a frame scan be allowed to assume the normal frame-scan time, T_0, by reduction of the basic velocity, v_{max}, so that the whole scanning process is slowed down, the bandwidth will be compressed in proportion to a value f_c.

Fig. 5(*c*) illustrates the waveform $e(t)$ resampled according to this reduced bandwidth f_c, where $f_c/f_0 = T/T_0$.

As a check on the calculation, let $D_F = 0$, i.e. a blank picture; the factor becomes $1/n$, as would be expected, since the scanning velocity is n times as great. When $D_F = 1 \cdot 0$ there is also no gain, since this corresponds to a complete chequerboard pattern.

(3.2) Half-Tone Reproduction

The analysis carried out in Section 3.1 is applicable only to black-and-white, as opposed to half-tone, reproduction, since all changes in level were supposed to occur at full amplitude. In general, detail will occur at all levels of contrast, and we shall consider now a change or sequence of changes at picture-element spacing, with an amplitude which is a fraction, a, of the maximum amplitude, as shown in Fig. 7. As indicated in Section 3.1, such sequences of changes may set up frequencies outside the prescribed band. Let us therefore determine what change needs to be made to K, the slope of the working characteristic [Fig. 4(*b*)], in order that such elements would be accommodated. Since the value of $|de/dx|$ for picture-element detail at full amplitude is by definition unity, a may be substituted for $|de/dx|$ in eqn. (12), and eqn. (17) becomes

$$\frac{nv_0}{1 + aK} = v_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

whence
$$K = \frac{n-1}{a} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

and eqn. (20) becomes

$$\frac{T}{T_0} = \frac{1}{n}\left[1 + (n-1)\frac{D_F}{a}\right] \quad . \quad . \quad . \quad . \quad (23)$$

It is now apparent that if $a = D_F$ no compression is achieved. For large values of n the saving tends to a value of D_F/a; thus, for $D_F = 1/20$ a saving by a factor of 2 would mean that any rapid changes of less than one-tenth of the maximum amplitude would not be resolved in the restricted bandwidth f_c.

The implications of this analysis are of sufficient importance to warrant detailed consideration. It has been shown that a bandwidth reduction can be effected by the means discussed only at the expense of the dynamic contrast range. For a black-and-white picture (comprising only two quantum levels), the maximum compression is obtained; but, in general, the compression will be considerably reduced, depending upon the smallest amplitude change which it is required to resolve. The reason is clear from inspection of Fig. 6, which depicts a waveform for which $|de/dt|$ tends to a constant. For such a waveform the frequency of alternation is inversely proportional to the amplitude of the alternation. If all changes occur at a similar amplitude, a constant value of $|de/dt|$ implies also a constant frequency, which will correspond with the average frequency of alternation occurring in the uncompressed signal. Hence, in the limit, using a maximum velocity v_{max} approaching infinity (i.e. $n \to \infty$), the factor of compression is numerically equal to the factor D_F, which is the average slope [see eqn. (5)]. For rapid alternation at lower levels of amplitude the frequency will increase, and at an amplitude level numerically equal to D_F no compression will result.

In practice, however, since the maximum velocity of scan will not be infinity, the compressed waveform will tend to have constant slope only in regions where rapid changes are occurring. Thus, with reference to eqn. (13) $|de/dt|$ will tend to a constant if $K|de/dx| \gg 1$. (See Fig. 6.) If the ratio of the maximum/minimum scanning velocity, n, was, for example, 10, then a transition occupying 10 picture elements or more would not be compressed to have the maximum slope, and would be transmitted in the compressed bandwidth f_c in spite of having a reduced amplitude. The slower the rate of transition between two grey levels, the more nearly equal may these grey levels be. Thus, in the practical case previously quoted, in which a picture with a detail factor of 5% was scanned with a velocity ratio of 10, the calculated saving of 7 : 1 would not be restricted to a pen-and-ink drawing but would also apply to a picture with a continuous range of half-tones in large areas where the transitions are relatively slow, but a restricted range in small areas where the transitions are very rapid.

A practical experiment would be necessary in order to determine the extent to which compression could be tolerated on this basis, but for certain applications the intelligence conveyed by such a picture would no doubt be considerable. Furthermore, small-area detail could probably be preserved to a considerable extent by accentuating the high-frequency response of the output signal from the camera before differentiation. It is noteworthy that the accentuation of small-area contrast is fundamental to the mode of operation of certain types of camera tube, and the picture produced by them would probably suffer only slight distortion with full compression.

In conclusion, it may be said that this particular system of compression has traded bandwidth for contrast range; however, a statistical redistribution has been achieved in that sharp boundaries in the picture will be transmitted only if they are of relatively high contrast ratio, whilst the greater part of the picture containing transitions of longer time will be transmitted with far greater range of grey values. The picture detail has been more uniformly distributed over the picture area than is the case with conventional constant-velocity scanning systems.

(3.3) Practical Considerations

It is not proposed to discuss in any detail the design of apparatus which would conform to the block schematic shown in Fig. 4(*a*), but brief consideration will be given to some of the significant practical aspects of such an arrangement.

(3.3.1) The Waveform to be Transmitted to the Receiver.

The output signal from the amplifier, $e(t)$ in Fig. 4(*a*), is the waveform required to be transmitted, and this inherently contains all the required information. The essential requirements for reconstructing the picture at the receiver are that the scanning velocities be identical at the receiving and transmitting ends (apart from a time-delay) and that the correct amplitude function $e(t)$ be received. However, it is unnecessary to transmit both a scanning-velocity function and the amplitude function $e(t)$, since the former may be derived from the modulus of the time-differential of $e(t)$. [See eqn. (7).] The information carried by the two functions is in duplicate. Thus, $e(t)$ is transmitted and $|de/dt|$ is determined in the receiver and used to control the time-base scanning velocity. The basic velocity v_{max} and the constant K need to be known only approximately at the receiving end, since these will affect, respectively, the horizontal width of the picture and the degree of information decompression; both are assessable by eye.

Picture synchronization could be effected in a manner similar to that developed by Bedford and Puckle,[4] since the basic problem is identical to that encountered in their velocity-modulation system, i.e. constant-voltage scanning could be used for the line scan and a ratchet circuit for the frame scan. Synchronizing signals derived from the flyback of the line- and frame-scanning waveforms could be transmitted in the normal manner.

An alternative scheme would be to employ vertical and horizontal black synchronizing bands on the camera mosaic, which would be scanned and compressed in bandwidth in the same manner as the picture. Thus, whilst the bandwidth would be compressed the positional accuracy on the receiver screen would be unaffected.

(3.3.2) Intensity Correction.

It is apparent that velocity modulation of the receiving-tube scanning beam will inherently bring about an undesired component of intensity modulation. All rapid transitions of picture intensity will be reproduced with excessive brilliance as compared with slow transitions. In order to eliminate this undesired component it will be necessary to apply also the velocity-control signal to the grid or cathode of the receiving tube, in the appropriate sense, as an intensity-correction signal. This will impose a limit on the maximum value of v_{max}/v_0 which it will be possible to use, and a ratio of about 10 : 1 would probably represent an upper practical limit. It is worth noting that as a result of intensity correction the beam current, and hence the aperture loss, would be reduced in areas of fine detail where high resolution is important.

(3.3.3) Use of Constant Frame-Scan Time; Control of v_{max}.

Eqn. (20) has shown that the lowest flicker rate (frame period $T = T_{max}$) corresponds to the picture having the greatest detail, and that T_{max} is decided by adjustment of the value of the basic velocity v_{max}. Pictures containing less detail than this maximum will have a flicker rate which is, in general, unnecessarily fast. It may therefore be advantageous to lengthen all frame-scan times to be equal to T_{max}. This would provide a constant-frame-time, but variable-line-time, system and this may assist a further compression step to be described in Section 5.2.

Working on a frame-by-frame basis, the total picture-detail in successive frames may be measured (integrated) and fed back to the camera scanner to control v_{max}. This control would need to be conveyed to the receiver by some suitable low-frequency modulation.

The system principle which has been described above is suited to the transmission of television pictures, within moderately compressed bands, where the interest value of the programme material is the overriding factor. Furthermore, the system is particularly suited to the transmission of black-and-white pictures in much reduced bands as, for example, might be employed for facsimile reproduction of letters and telegrams.

(3.3.4) Response Time of Velocity Control.

The analysis given in Section 3.1 tacitly assumes that the action of the velocity-control circuit is instantaneous. In practice, of course, it will depend upon the bandwidth of the closed feedback loop, and this must be made large enough to accommodate the spectrum which will result when a boundary is scanned at the maximum velocity of scan, v_{max}. For example, consider a sudden transition from black to white occurring at some arbitrary position along a scanning line; the scanning beam will approach the transition with a velocity n times as great as that which would be required to resolve the transition in the reduced bandwidth f_c. If the bandwidth of the feedback loop was limited to f_c the scanning beam would overshoot the transition by n picture elements before the instruction to slow down was acted upon. However, if the bandwidth of the feedback loop is made equal to nf_c the position of the transition will be observed with an accuracy of one picture element. A filter of cut-off frequency f_c may, however, be inserted in the transmission circuit, since by virtue of the variable-velocity scan the transition will have been extended in time and will therefore be reproduced with the same positional accuracy, in spite of the reduced bandwidth. In practice, the bandwidth-limiting filter is bound to introduce some degree of distortion, but provided its phase characteristic is linear, or substantially so, positional errors will not arise.

(4) VELOCITY MODULATION APPLIED TO FREQUENCY COMPRESSION

The basic shortcoming of the system just described arises because of its unsuitability for *frequency* compression. There would be a gain only if large jumps in brightness level were less probable than small jumps, which is a doubtful assumption; let us consider the principles to be somewhat modified.

(4.1) Information in a Half-Tone Picture

The essential detail in an amplitude-quantized half-tone picture has already been considered in Section 2.2, and only a brief reference will be made here. Fig. 2(*a*) shows a typical picture waveform and Fig. 2(*b*) the quantized form of it. On the basis which we are adopting there are two kinds of symbol to be transmitted:

(*a*) Adjacent samples which have identical quantum levels (which define areas of uniform greyness). These have high probability, $p_i(j)$.
(*b*) Adjacent samples which have different levels (which define all edges and boundaries of the picture). These have lower probability, $p_i(j)$.

Photographic examples of such quantized pictures are to be found in Reference 5. They consist of flat grey surfaces and of bounding edges; it is required to abbreviate the former and stretch the latter to equalize more nearly $p_i(j)$.

(4.2) Frequency Spectrum of a Picture Edge (Transition)

At constant scanning speed all transitions between any two quantum levels take place over a sample time $\Delta t = 1/2f_0$. If the picture is stored on a mosaic, we consider instead a sample distance Δx; the sample time and the corresponding spectrum bandwidth f_0 then depend upon the scanning velocity. Fig. 8 shows some typical transitions; in Fig. 8(*a*) two typical jumps

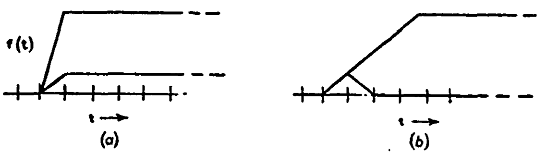

Fig. 8.—Typical transition regions.
(*a*) Equal frequencies.
(*b*) Equal slopes.

between different steady-state grey levels are shown; the amplitudes, and hence the slopes, are very different, but they require the same frequency bandwidth f_0. On the other hand, Fig. 8(*b*) shows two transitions of identical slope, but one continues on with the same slope to a higher level, whilst the other reverses its slope to the original level. Then the first has a smaller bandwidth than the second.

Clearly, the bandwidth required for any complete transition is inversely proportional to the time taken to pass between consecutive steady-state levels. Fig. 9(*a*) illustrates a typical wave-

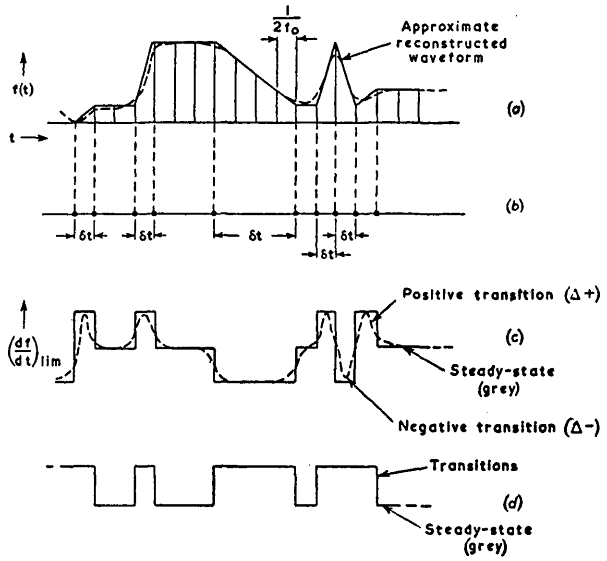

Fig. 9.—Transition and steady-state regions of a quantized picture waveform.
(*a*) Typical waveform.
(*b*) The intervals δt.
(*c*) Differentiated (limited) waveform.
(*d*) Modulus differential.

form containing different transitions; these transition times are marked δt in Fig. 9(*b*). Such times are quite independent of the quantum magnitudes of the transitions. (In this Figure the approximate reconstructed waveform is shown in idealized form which would result from passing the pulse samples through a filter of bandwidth f_0.) Let us refer to these positive and negative transitions as ($\Delta+$) and ($\Delta-$), whatever their magnitudes.

(4.3) Velocity Modulation for Equalizing Transition-Spectra Bandwidths

It is required to recode the signal waveform, to achieve a frequency compression, by velocity modulation of the scanning beam. Fig. 9(*c*) shows the waveform $f(t)$ differentiated and the resulting waveform amplitude limited, so as to produce a waveform having two possible equal positive and negative amplitudes. Call this $(df/dt)_{lim}$. The dotted curve illustrates the resulting waveform if a filter of bandwidth f_0 is used. This separates out the three regions, namely those of any positive transition $(\Delta+)$, those of zero transition (steady-state) and those of any negative transition $(\Delta-)$. Ideally, the compression may now be considered in two stages:

(*a*) Near-equalization of the total steady-state time and the total transition time (+ and −) over some suitably long period. Most conveniently this would be a line-scan period. Fig. 9(*d*) shows the modulus differential $|df/dt|$.

(*b*) Near-equalization of all transition periods δt. This equalizes the bandwidth required for all transitions.

These processes may be too complex to achieve in practice, and a compromise should perhaps be considered. This might consist in velocity modulation of the scanning beam so as to near-equalize all successive intervals, both steady-state and transition, i.e. to near-equalize the successive intervals between all the points (marked with dots) on the time-axis in Fig. 9(*b*). Such a process amounts to treating as equally important all the edges and boundaries of the quantized picture, as considered in Section 4.1.

(4.4) Instrumentation of Transition-Time Equalization

In the original scheme, as outlined in Section 3, the compression of the picture was achieved by a closed-cycle scanning system in which waveform slope was fed back, so that the response curve of the system tended to have a constant slope (Fig. 6). Such a scheme is suited more to the compression of picture slopes than picture frequencies.

For compression of frequencies it is preferable to use a closed-cycle scanning system in which frequency is fed back. As we have seen in Section 4.3, the frequency of a transition depends on the transition time between successive steady-state levels, i.e. upon the spacing between the consecutive zero points in the differential coefficient of the picture waveform. This has already been discussed and illustrated by Fig. 9.

It is essential to consider the picture as quantized, not continuous. This may be accomplished by a preliminary high-speed scan of the camera image at a linear and very high velocity, followed by quantization and transference on to storage mosaic, as shown in the block-schematic of Fig. 10(*a*).

When scanned, the quantized picture results in a step-like waveform, all the step edges of which are equal in time duration, i.e. a constant transition time. This waveform is differentiated to produce a series of equal-duration pulses with amplitudes corresponding with the transition amplitudes. Severe amplitude limitation, if applied, will result in a series of pulses of constant duration and amplitude. The duration of the pulses will be determined by the bandwidth, f_0, of the filter interposed between the differentiator and the limiter, and this must be such that the two adjacent pulses never overlap, even when they are one picture-point apart. These identical pulses define where the quantized picture transitions occur [corresponding to the dots in Fig. 9(*b*)] and so may be used to control the speed of the time-base used for scanning the stored quantized picture.

An electrical switch triggered from the pulses will serve to make the time-base travel slowly or rapidly according to whether a pulse is present or not. Thus, the control waveform is divided into steady-state horizontal regions and transition regions. The scanning-beam velocity is thus made high for the first regions and low for the second.

Such a scheme recognizes only two velocities, fast and slow, whilst the number of distinct grey (quantum) levels dealt with is, say, ν. This is in strict contrast to the compression scheme dealt with in Section 3, in which, in general, ν velocities can be used with only two quantum levels (black-and-white pictures).

The waveform transmitted to the receivers is the compressed picture waveform $e(t)$, the output from the storage mosaic. At a receiver the processes of differentiating and limiting may be reproduced, so that the scanning-beam velocity follows that at the transmitter. Fig. 10(*b*) shows a block schematic. Very little increase in complexity of technique will be required. Most of the difficulty involved in frequency compression is laid on the transmitter; this is a valuable attribute of such a principle of compression.

(5) COMPRESSION IN THE VERTICAL DIMENSION: CORRELATION BETWEEN ADJACENT LINE SCANS

So far, only compression in the horizontal direction, i.e. in the direction of normal scanning, has been considered, but similar potentialities exist for compression in the vertical direction. Redundancy arises by virtue of the fact that, given the form of a television scene up to say the nth scan line, the $(n+1)$th line may be guessed to a high degree of accuracy. More precisely, there exists correlation between adjacent lines, or—by an

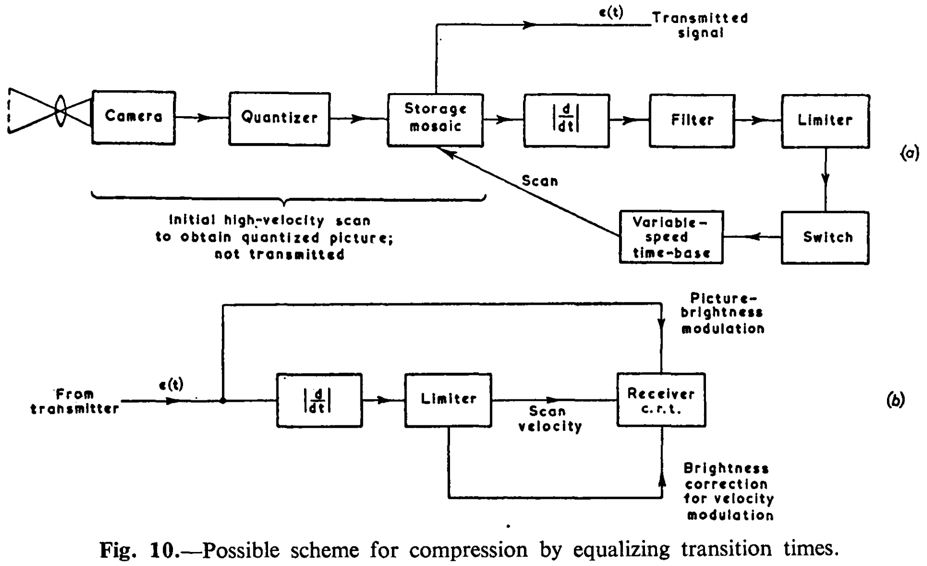

Fig. 10.—Possible scheme for compression by equalizing transition times.
(*a*) Transmitter. (*b*) Receiver.

alternative point of view—the transition probabilities between successive line-scans are substantial only for similar scans. This point of view is similar to that adopted in the preceding Sections, except that we are now considering line-scan waveforms as the successive symbols rather than picture points or sample ordinates. Before continuing with suggestions for reducing this redundancy, let us digress for a moment to stress a rather general aspect of communications.

(5.1) Recognition of Form; Visual and Aural Patterns

In an earlier paper[2] one of the authors has emphasized the relevance of a branch of modern psychology to the communication engineer, and the practical advantages of a common understanding with psychologists in this field. This branch is a very physical one, and concerns *Gestalt*, or the recognition of form. It concerns questions such as: how do we recognize a person's face, from all distances and most angles of view, or recognize an object of simple geometric shape mixed with random shapes? The subject is relevant to television in that this essentially involves the recognition of patterns which may be mixed with noise or other disturbing influences. Very similar principles apply to sound. In normal conversation we recognize a friend's voice, although the background may be noisy and disturbing and may indeed contain the sounds of other voices. What are the elements of recognition in speech? The problem is perhaps the most basic in telephony; the listener is attempting continually to recognize the speaker's words and phrases or the emotional stressing of them. The future development of telecommunication depends to a large extent on an understanding of these elements of recognition, because if they can be determined they may be employed for the coding of the messages to overcome noise most effectively, yet to eliminate as much redundancy as possible.

When viewing a scene, the eye scans its centre of sharp vision (a narrow angle of only a degree or so) over the scene, not in the strict television sense of parallel lines but along some curve which depends upon the scene itself. It most probably travels quickly over those parts of the scene which the observer has already seen (i.e. the steady-state background) but spends longer over new, changing or moving parts, since, being less probable, these contain more information. This is pure conjecture, but—it will probably be agreed—fairly reasonable conjecture. As the eye scans the scene (or pattern) a correlation will be found between successive parts of the scene which convey the idea of its form, and, indeed, distinguish pattern from purely random noise.

However, such suggestions may appear to be rather vague; consider instead a television picture with its precise geometrical form of scanning. What distinguishes a pattern from random noise is partly its predictability by virtue of its continuity of outline; such continuity implies that, given the pattern shape up to, say, the nth line, the waveform of the $n(+1)$th line-scan may be predicted to a high accuracy. The correlation between successive line scans is close, i.e. $p_n(n+1)$, the transition probability distribution for the $(n+1)$th line given the nth, is narrow.

Thus, two methods suggest themselves for compressing the television-signal information by reducing the redundancy, yet retaining the essential elements of the picture. These depend respectively upon the correlation existing between successive lines and the transition probabilities between adjacent picture-points (i.e. sampled ordinates) in the vertical Y-direction. These two principles are really similar, but the techniques involved differ.

(5.2) Potentialities for Compressing Picture Information in both Vertical and Horizontal Scanning

(5.2.1) Correlation between Adjacent Lines.

It may be argued that, owing to the high degree of correlation between successive line-scans, only their difference need be transmitted, since only this represents new information per line. It may therefore be possible to send the first line-scan (top of picture) at a very low speed (e.g. taking a whole frame-scan period) in order to keep the frequency band narrow, and then to transmit only successive line-scan difference or error signals. Using the variable-velocity-scan principle this will achieve a compression by virtue of the reduced picture-detail to be transmitted [see eqn. (4)].

Such a technique would probably involve great difficulties; for one thing the successive difference or error signals would need to be accumulated from the top line to the bottom and the precision of reassembling the picture might vary over the frame.

Delbord[7] has suggested a method which might be developed for correlating successive frames. The method he suggests would reduce considerably the amount of information to be transmitted, but the grouping of transitions would still require to be equalized if bandwidth was to be saved (see Sections 4.3 and 4.4).

(5.2.2) Transition Probabilities in the Vertical Y-Direction.

It may prove possible to apply precisely the same velocity-scanning principle in the vertical Y-direction as already described for the horizontal X-direction. The picture is scanned twice, alternately in X- and Y-directions.

At the transmitter the camera mosaic may be scanned in the X-direction and each line of one complete frame compressed, as illustrated by Fig. 5. Instead of being transmitted, however, this frame may be stored upon a second mosaic (by a subsidiary constant-velocity scanning process) which may then be scanned in the Y-direction by an identical closed-loop velocity-scanning circuit (Fig. 10). The output waveform is transmitted and contains all the information for the receivers. While this Y-scanning of the second mosaic is taking place, the first mosaic, in the camera, lies dormant and unscanned.

In the receivers the waveform is stored on a subsidiary mosaic and then rescanned and presented on the viewing screen. These two stages are similar to the two stages in the transmitter. First, the received waveform is written on to the storage mosaic by a velocity-controlled Y-scanner; secondly, this mosaic is scanned in the X-direction by a velocity-controlled scanner and the resulting signal presented on the viewing screen as the reassembled picture. Note that the received picture is two frames delayed on the transmitted picture and that alternate frames are empty. Fig. 11 illustrates the scanning succession. Thus, a time-wasting factor of 2 : 1 is involved. However, such a system

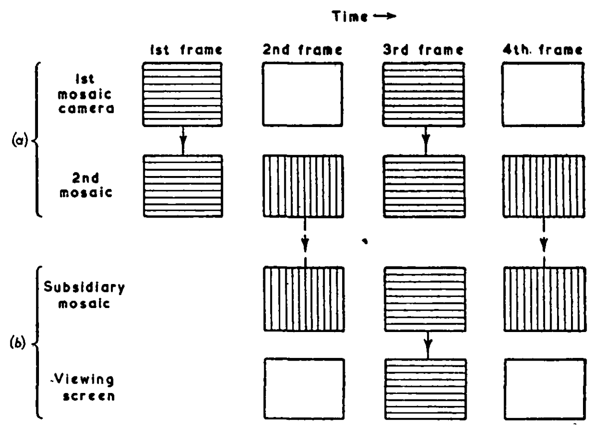

Fig. 11.—Compression in two dimensions.
Alternate X- and Y-velocity scanning, with storage mosaics.
(a) Transmitter.
(b) Receiver.

might readily be suited to colour television, on a basis of successive coloured frames.

Such a system depends for its action upon successive frames occupying identical scanning times, T_s, in order that one frame may be stored during the same interval that the preceding frame is being read off. This is possible if the method of controlling the basic scanning velocity, v_{max}, is used, as described in Section 3.

It might appear that this principle would result in the possibility that the bandwidth-compression ratio could be squared; this is not so, however. Note that after compression in the X-direction, at the transmitter, the waveform must be stored upon a second mosaic using a constant scanning velocity, otherwise the picture would be decompressed. Consequently, different lines on this mosaic will occupy different lengths, according to the amounts of detail contained therein. However, these line lengths will not be random, but will probably vary smoothly from line to line, owing to the correlation between successive lines, as has been discussed in Section 3.1.

Thus, when the second scanning process is applied, in the Y-direction, the mosaic being scanned will contain large areas containing no detail whatever. It may be possible to restrict these areas by not scanning over the entire width of the storage mosaic; this would be possible only on the assumption that no one line in the X-direction ever contains the full detail of a chequerboard.

Such proposals obviously require a considerably more complex technique than the present-day wide-band systems, but this is the price to be paid for the great bandwidth compression which is potentially available for black-and-white or colour systems. However, it is thought that the type of technique involved is quite practicable nowadays or in the not too distant future. Techniques of mosaic storage, closed-loop control and, if required, amplitude quantization, have recently undergone extensive development, especially in connection with high-speed digital computors.

(6) CONCLUSIONS

A method of bandwidth compression has been described which is applicable in principle to various systems of communication, especially telephony and television. The potential compression of television signals has been shown to be very great in both X- and Y-directions. Whether or not such a high degree of compression would ever be required, or its attainment attempted, is another question, which is largely an economic one.

So far, no mention of noise level has been made; this omission has been deliberate. The principle of compression, dependent upon coding and the resulting change of the statistical structure of television signals, is independent of noise. But the degree of compression attainable in practice, and the reduction of signal redundancy, depend entirely upon the noise level and the tolerable distortion of the reconstructed picture. This is one of the advantages of considering the compression problem from the point of view of communication theory. The ultimate degree of compression attainable in the presence of a specified noise level, and methods of coding to achieve this limit, are not considered in the paper.

The type of "recoding" considered here is for the reduction of excessive redundancy, and of course is not noise-combating; it is most likely to be applicable to systems other than entertainment television.

(7) ACKNOWLEDGMENT

The authors wish to thank Prof. Willis Jackson, Head of the Department of Electrical Engineering, Imperial College of Science and Technology, and the Chief Engineer of the B.B.C. for permission to publish the paper.

(8) REFERENCES

(1) SHANNON, C. E.: "A Mathematical Theory of Communication" (University of Illinois Press, Urbana, 1949).
(2) CHERRY, E. COLIN: "A History of the Theory of Information," *Proceedings I.E.E.*, 1951, **98**, Part III, p. 383.
(3) SHANNON, C. E.: "Prediction and Entropy of Printed English," *Bell System Technical Journal*, 1951, **30**, p. 50.
(4) BEDFORD, L. H., and PUCKLE, O. S.: "A Velocity-Modulation Television System," *Journal I.E.E.*, 1934, **75**, p. 63.
(5) GOODALL, W. M.: "Television by Pulse Code Modulation," *Bell System Technical Journal*, 1951, **30**, p. 33.
(6) GOURIET, G. G.: "A Method of Measuring Television Picture Detail," *Electronic Engineering*, 1952, **24**, p. 306.
(7) DELBORD, Y.: "La Cybernétique." Réunions d'études et de mises au point. (Published by "Editions de la Revue d'Optique Théorique et Instrumentale," Paris, 1951), p. 207.

(9) APPENDIX

Referring back to Section 2.2, it has been considered advisable to discuss compression in terms of a picture-detail function, D [see eqn. (4)], rather than use Shannon's information formula, eqn. (2), because D is more directly related to the mechanics of the system of compression described. It is intuitively clear that, over a wide range of possible forms for the distribution $p_i(j)$ the factors D and H might be approximately equal in a quantitative sense. Of course, the actual form of $p_i(j)$ can be determined only by experiment.

However, it has been pointed out by Gabor that there is one particular distribution function for which D and H are identical and which it would not be unreasonable to assume as practice.

Shannon's definition of the rate of information is:

$$H = -\sum_k p_k \log p_k$$

Considered as a succession of pairs of symbols, the first with a probability p_i followed by a second with a transition probability $p_i(j)$, this may be written in accordance with eqn. (2) as

$$H = -\sum_i \sum_j p_i p_i(j) \log p_i(j)$$

Let i and j now stand for amplitudes of the television-signal successive samples; the picture detail per sample [see eqn. (4)] may be written:

$$D_l = \sum \sum p_i p_i(j) |j - i|$$

If now we assume the distribution

$$p_i(j) = \exp - k^2 |j - i|$$

then $$H \propto k^2 \sum_i \sum_j p_i p_i(j) |j - i|$$

which is proportional to the picture detail D_l.