A Companded One-Bit Coder for Television Transmission

By R. H. BOSWORTH and J. C. CANDY

(Manuscript received November 19, 1968)

We compand a one-bit coder by increasing its step size when a string of equal bits is detected in the transmitted code. To code and decode each string we use a weight sequence 1, 1, 2, 3, 5 $\cdots$ 5, the weight returns to unity when the string ends. Stability considerations restrict the choice of weights but those proposed give adequate stability as well as improve the signal-to-noise ratio about 5 dB. The weighted coder has a wide tolerance to changes of input, so that a ±3 dB change from the design value is hardly visible to most observers. Matching weights at the transmitter and receiver is uncritical because mismatches appear as small changes of contrast rather than as noise. The circuit is easily implemented because it is tolerant to changes of component values.

There is a description of an experimental coder and decoder, together with subjective and objective measures of performance. Signal-to-noise ratios of 50 dB are reported.

I. INTRODUCTION

Encoding an analog signal to digital form entails quantization of amplitude. This process introduces a noise into the analog signal that is recovered from the digits. The magnitude of the noise, relative to the signal, is determined by the bit rate in the digital representation and the spectrum of the signal. Successful coder designs make efficient use of the digits, avoiding worthless redundancies, and shape the noise to be subjectively least noticeable.

Delta modulation is one of the simplest and best known coding methods.[1] It changes its analog output positively or negatively by a fixed increment at regular instants, as illustrated by V in Fig. 3. Differential coding is a related method where, at regular instants, the output changes by any one of a set of prescribed values. Delta modulation is regarded as one-bit differential coding because at each sampling time it transmits either of two codes, a pulse or a space, representing a positive or a negative step, respectively. In general, an m-bit coder transmits one of 2^m codes at each sample time.

The advantages of one-bit coding are simplicity of circuitry and a high sampling rate. Thus, for a given bit rate in the digit channel, its sampling rate is m-times greater than that of a corresponding m-bit coder. Although the total noise power from a one-bit coder is greater than that from a multibit coder, much of the power occurs at higher frequencies where it is out of the signal band. The advantage of multibit coding is the ability to grade the step sizes to suit the signal values.[2] Thus, some large steps are provided to track large changes and some small steps are provided to accurately reproduce fine details. In this respect, ordinary one-bit coders are handicapped by having only a single step size.

Theoretical results by J. B. O'Neal show that multibit differential coders have larger signal-to-noise ratios than delta modulators.[3] Practical measurements confirm this and show that much of the advantage comes from companding the quantization levels.* We describe a method for varying the step size of a one-bit coder which has the advantages of both companding and a high sampling rate.

Several authors have described a method for companding delta modulators by changing the step size according to the average pulse rate in the digit channel.[4–7] The steps are smallest when there is an equal number of pulses and spaces; they increase when there is a higher proportion of either pulses or spaces for a significant time. This technique has been used for audio signals to adjust the step size with loudness and pitch of the sound.

For video signals we usually are directly interested in the time dependence of the signal and so require means for adjusting the step size according to instantaneous signal values rather than an average value. Suitable methods have been described by M. R. Winkler and J. E. Abate.[7,8] They vary the step size when certain pulse patterns are detected in the digit channel. Thus, steps are increased when a string of consecutive pulses or spaces are detected. This paper describes the design, construction, and performance of such a coder. It differs from earlier coders in the way step sizes increase and decrease and in that the companding is incorporated in a direct feedback coder instead of a delta modulator. Direct feedback coding, which is reviewed in Section II, is an improvement on differential coding.

* Two-bit coders have insufficient levels to permit adequate companding so they usually are inferior to other coders.

II. DIRECTING FEEDBACK CODING

Direct feedback coders are described in Ref. 9. They function almost the same way as differential coders, but the circuit is arranged to allow greater flexibility of filter design. Figure 1 is a block diagram of a one-bit direct feedback coder and Fig. 2 shows some typical filter characteristics. For television signals the de-emphasis filter H_2 is a short time integrator; the pre-emphasis H_1 is a differentiating filter approximately the inverse of H_2; and the filter A in the feedback loop is a long time integrator.

The feedback acts like a servomechanism trying to make the average value of the quantized signal y equal to the pre-emphasized input x. The difference between x and y is accumulated in A and used to correct the quantized output. The quantized signal in a one-bit coder is observed to oscillate between a positive and a negative level in such a way that its average equals x, as Fig. 3 demonstrates. Changing the pattern of oscillation, the coder interpolates values between the quantization levels, but low frequency components of the oscillation appear as granular noise on the output. The filter A is chosen to make these low frequency components small. High frequency components are de-emphasized by the integrating filter H_2 whose output steps up or down in response to a pulse or a space as does the output of a delta modulator. The advantage of direct feedback coding is flexibility in choosing the deemphasis H_2 independent of the interpolation process which is controlled by the feedback loop.

Notice in Fig. 3 how the large voltage spike in x overloads the quantizer by exceeding its quantization level. The coder responds with a string of pulses which is the largest signal it can transmit. The resulting distortion of the signal is called slope overload; it is a

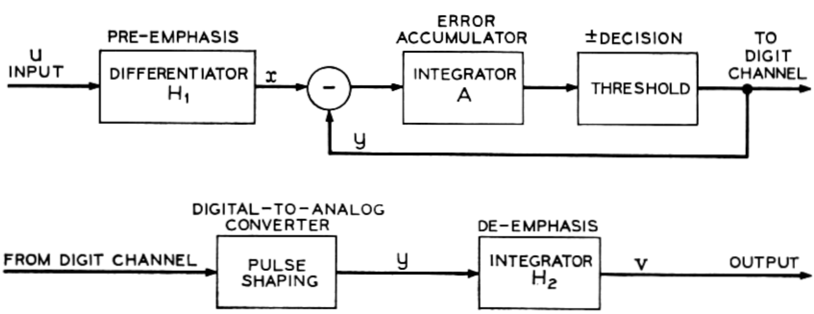

Fig. 1 — A one-bit feedback codec.

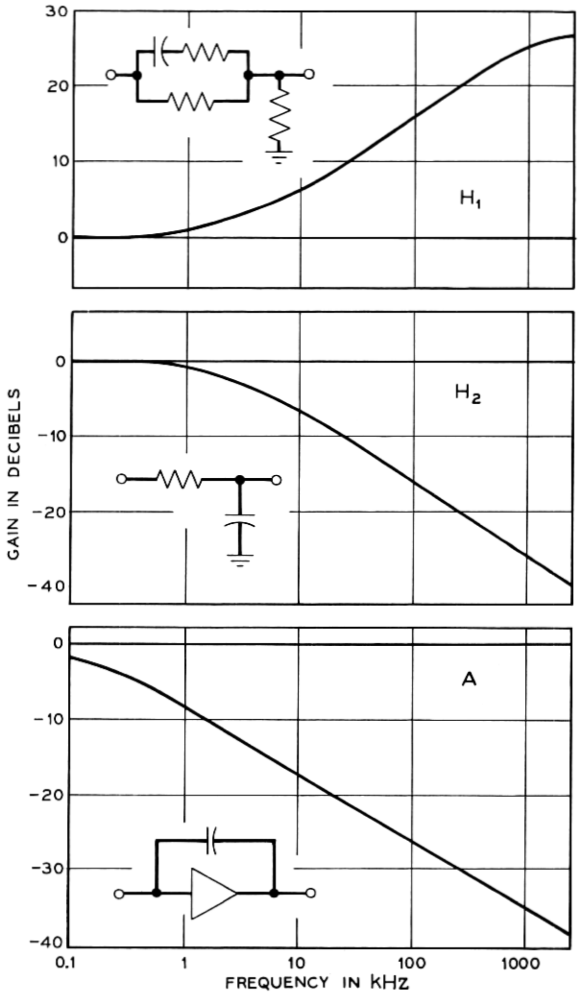

Fig. 2 — Filter characteristics.

characteristic of systems using integrating deemphasis. Usually, the input to the coder is adjusted to a compromise where there is neither too much slope overload at edges nor too much granular noise on "flat" areas.

In the companded coder, overloading is detected by locating strings of pulses or spaces in the code, and then the step size is increased both at the transmitter and at the receiver. This increase extends the range of the coder but increases the interpolation noise in the vicinity of sharp changes. The success of the technique depends on the way step sizes are varied. We have no theoretical criterion for optimizing the formula, instead practical reasons are given in favor of the proposed scheme. The strongest arguments are: the system works, it is easy to implement, and it functions as well as any scheme we have tried for the *Picturephone*® see while you talk service, which is approximately the transmission of a 1 MHz video signal as a 6 MHz binary signal.

III. COMPANDING METHOD

3.1 *The Weights*

Figure 4 shows a block diagram of the proposed coder and decoder (codec). It differs from the ordinary feedback coder by the addition of a weighting circuit in the feedback path at the transmitter and in series with the receiver. The weighting is controlled by a circuit that detects strings of pulses or spaces in the transmitted code. The signal y is then made up of a pulse sequence whose amplitudes depend on the code. The pulses corresponding to the first two bits of each string are left unweighted at the smallest step size. For the third and fourth bits the pulse size is increased two and three times, respectively. For the fifth bit, and all that follow in the string, the pulse size is made five times that of the smallest pulse's value. The string ends when a change of polarity is called for by the appearance of the complementary binary code; then the weight returns to unity. An example of a digit stream and its corresponding quantized signal is given in Fig. 5 which also shows the decoded signal. Compared with Fig. 3 there is a decided improvement in the reproduction of the signal because slowly changing signals are reproduced with smaller steps but the larger signal changes are reproduced with larger steps. Consider the reasons for using this particular set of weights.

3.2 *Choice of Weights*

The plan is to increase the step size when the input changes rapidly. Thus, small steps are used when the differentiated input x is small, and they are increased as x increases. In this way we take advantage of the fact that noise in busy areas of a scene is less noticeable than noise in flat areas.

The step size is left unchanged at its smallest value when no more than two consecutive bits are the same. Such codes are used to transmit

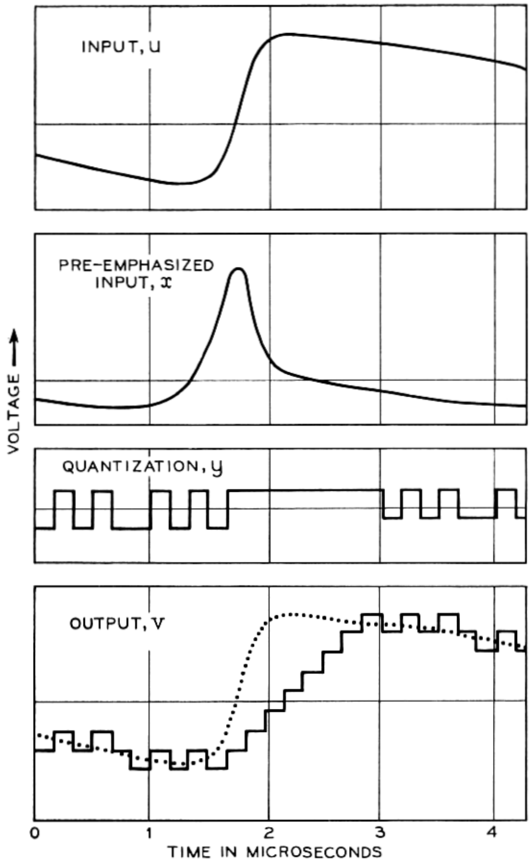

Fig. 3 — Waveforms in the ordinary one-bit direct feedback codec.

the slowly varying inputs that represent the flat areas of a picture. The largest of these codes is 110110 · · · 110 and the smallest is 001001 · · · 001. They correspond, respectively, to values of y whose average is $+\frac{1}{3}$ and $-\frac{1}{3}$ of the smallest step size. These codes are generated when x has a steady value in the range $\pm\frac{1}{3}$ of a step.

When x exceeds $1/3$ of a step size, codes with more than two repeated bits are generated; the weighting circuit then increases the step size. The signal level in the coder is set so that this occurs only in busy areas of the picture and at edges. Usually, the larger values of x appear as spikes of voltage resembling the one in Fig. 3. Therefore, step sizes should be increased promptly in order to code the transient in a short time; they should be promptly decreased afterwards. Indeed, it is desirable to code sharp changes in video signals in less time than is used to scan three picture elements, otherwise the distortion is objectionable.[9] For *Picturephone®* visual telephone, the edges should be coded in less than 1.5 microseconds, that is, with less than nine bits.

A stability requirement restricts the way weights can be applied. Consider for example, a poor design using a weight sequence 1, 1, 2, 4, 9 · · · 9 to code each string of similar bits. Figure 6 shows an impulse in the voltage x and the subsequent behavior of the quantized signal y: it oscillates continuously between the largest weights after the impulse instead of falling to unity. This oscillation is undesirable because it may increase granular noise in the flat areas of the picture. Figure 6 also shows the response of the proposed coder to an impulse. There is a small undershoot following the representation of the impulse but the step size assumes its smallest value after taking eight bits to code it.

A condition for the weights to fall to their lowest value after any impulse in x is that the weight sequence increase no faster than 1, 1, 2, 4, 8 · · · that is, each weight be no greater than the sum of previous weights in the sequence. The proposed weights 1, 1, 2, 3, 5 · · · 5 satisfy this requirement, giving a safe margin to dampen oscillations.

The weight returns to unity when a string of similar bits end. Then subsequent weight values are independent of the previous code which

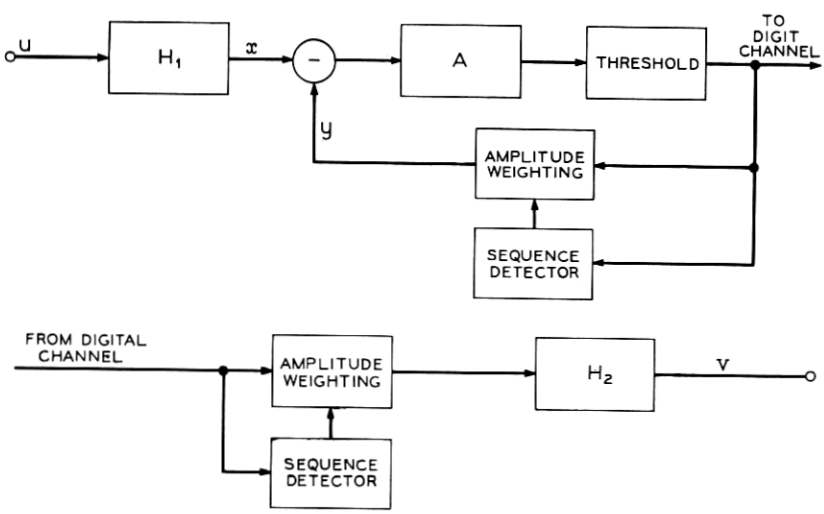

Fig. 4 — The companded codec.

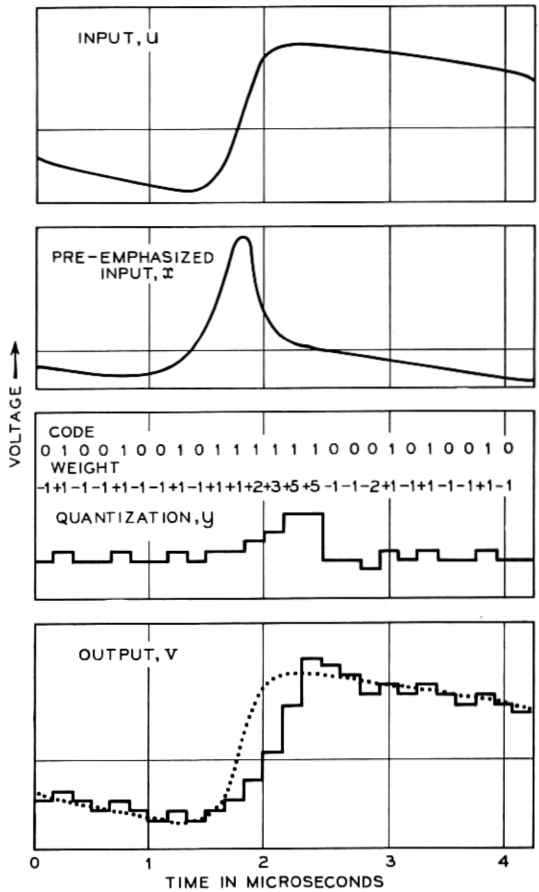

Fig. 5 — Waveforms in the companded one-bit codec.

helps to reduce streaking caused by transmission errors. Properties of a coder are often as dependent on its method of construction as they are on the philosophy of its design. In order that the evaluations be meaningful the circuits are described in the appendix.

IV. EVALUATION OF THE CODER

4.1 *The Test Setup*

For the tests the coder uses a 6.3 MHz sampling rate to code a television signal having 1 MHz bandwidth. This signal represents a 271 line interlaced picture, displaying 30 frames a second. All subjective

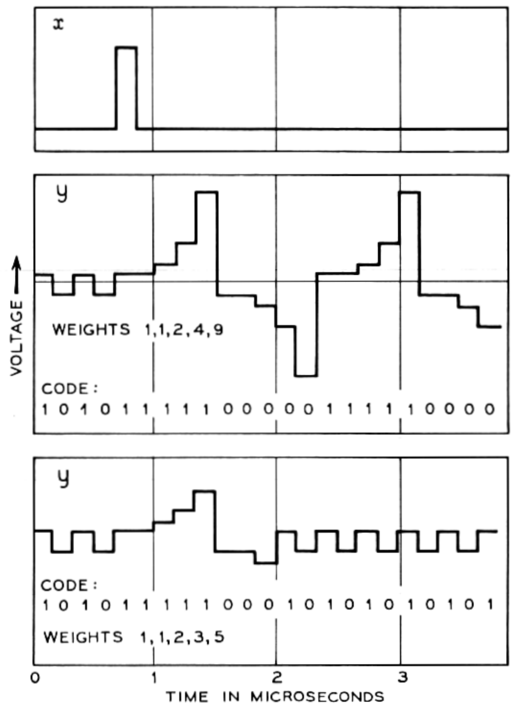

Fig. 6 — Responses to an impulse of x using two different weighting sequences. In normal use the signal will be band-limited so the impulse will be broadened.

tests were carried out using a 5½ by 5 inch display viewed from 3½ feet. The peak luminance was 70 foot lamberts and the room illumination about 100 foot candles.

4.2 *Subjective Tests*

Subjective tests were made by observers who were experienced in picture evaluation and familiar with the coding process. They compared two displays which they switched alternately onto the monitor with equal contrasts. One was the coded picture, the other an uncoded picture with noise added. Each observer varied the noise amplitude until the displays had equal overall quality for him. At this setting the ratio of the signal to the added noise power was recorded as his measure of picture quality. The noise used in these experiments was approximately gaussian with a flat spectrum from 100 Hz to 0.6 MHz as shown in Fig. 7.

The first group of tests concern the signal level in the coder. The

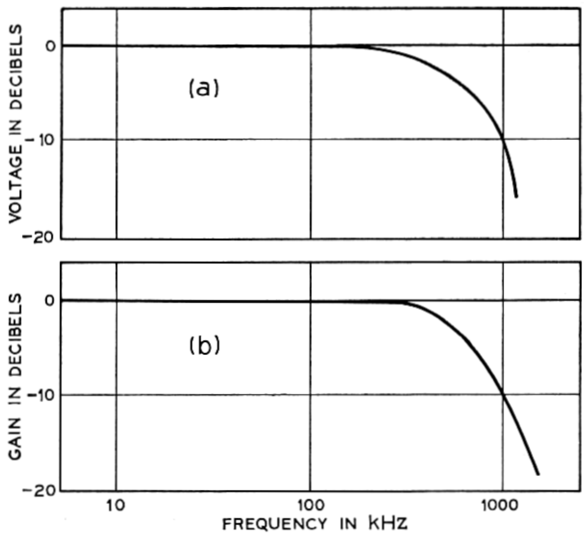

Fig. 7 — (a) Noise source used for testing; (b) Characteristics of filters used for restricting the signal band at the codec input and output.

input amplitude to the coder was varied while a compensating variation at the output maintained a fixed contrast on the monitor. At each amplitude setting an equivalent signal-to-noise ratio was obtained as described. Figure 8 shows the graph of signal-to-noise ratio plotted against signal amplitude into the coder. Results obtained by four observers are given.

Observers agree with one another for small inputs but differ at larger amplitudes where overloading predominates. They all prefer inputs around 70 mV; above this value the quality of the picture falls abruptly because overloading becomes objectional at edges of the scene. When the amplitude of the signal is decreased from 90 mV to 30 mV the picture quality falls slightly as there is a subtle interchange between overloading and granularity. Below 30 mV the granular noise becomes objectional. This graph was obtained using a video signal derived from a back lighted transparency that has unnaturally high contrasts. Figure 9 is a print of this film.

Figure 10 shows an evaluation of a natural live subject. This test was difficult to perform because of the high quality of the coded picture. Observers accept larger inputs (up to 120 mV) because movement makes edge distortion less noticeable. Figure 11 is an evaluation of a transparent resolution chart. This has the lowest signal-to-noise ratio because the peak to root-mean-square value is small, and because a rapid succession of black and white vertical stripes induced oscillation in the weights. But these patterns are unlikely to occur in real scenes: the codes usually transmit pictures of graphic material with little impairment. All of the graphs demonstrate the wide tolerance of the coder to changes of input amplitude.

A result of the second group of tests is shown in Fig. 12 demonstrating the benefits of weighting step sizes. Curve (a) in Fig. 12 is a subjective measure of the ordinary unweighted one-bit coder; the other curves are for the weight values specified on the graph. Notice that at low signal amplitude, where granular noise predominates, the weight has no effect. Weighting only improves the response to large inputs where overloading is important.

The next test concerns the tolerance of the coder to changes of weight values. The sequence 1, 1, 2, 3, 5 ··· 5 was proposed for our application; an attempt was made to find a better sequence experimentally. Figure 13 compares the proposed weights with the best we could find; there is little difference. In fact, the choice of weights is not critical provided they do not cause instability.

The last subjective test concerns the matching of weights at the transmitter and the receiver. Figure 14 shows the equivalent signal-

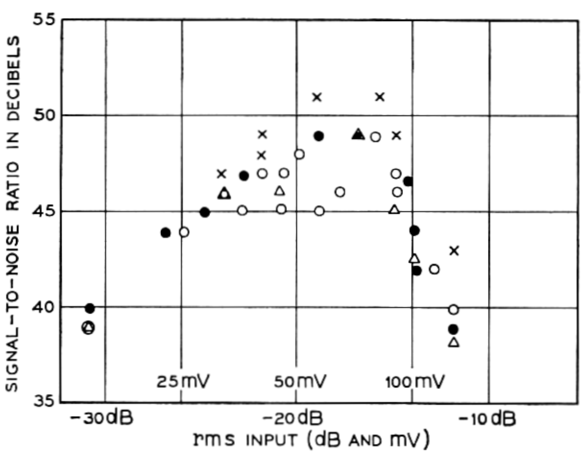

Fig. 8 — A ratio of peak-signal to root mean square-noise plotted against input amplitude. The noise was measured subjectively by four observers. The scene was a photograph of a face.

Fig. 9 — The still picture used for the subjective test in Fig. 8.

to-noise ratio of a coder with a 70 mV input and a weighting sequence 1, 1, 2, 3, 5 $\cdots$ 5, at the transmitter. At the receiver the weighting sequence was*

$$1, 1, (2 + \epsilon), (3 + \epsilon), (5 + \epsilon) \cdots (5 + \epsilon)$$

where ϵ is a controlled variable: it is the absissa of the graph. Graphs for other weight sequences are also given. In all cases the circuit is unusually tolerant of mismatching the transmitter and receiver. Mismatching weights tends to distort the scene in busy areas, rather than introduce noise. This is discussed in the Section A.2 of the appendix.

We have not obtained numerical evaluation of the effect of transmission error. The opinion of most observers is that error probability

* This type of mismatch is consistent with the method of construction where each new weight value is obtained by augmenting the previous one, as shown in Fig. 22.

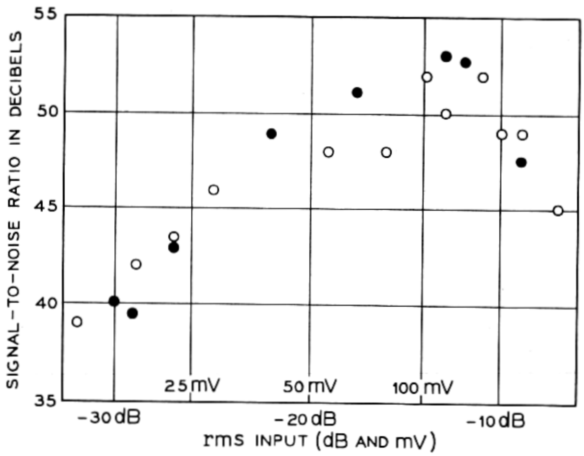

Fig. 10 — Subjective signal-to-noise ratios using a live scene.

less than one in 10^6 is hardly noticeable in a live scene. Errors more frequent than one in 10^5 were troublesome, but the picture was useful with error rates up to one in 10^3. Each error appears as a streak no longer than 0.6 inches with random amplitude. Synchronizing errors were not included because the timing signals were sent on a separate channel.

This subjective measurement is a valuable tool in that it gives more

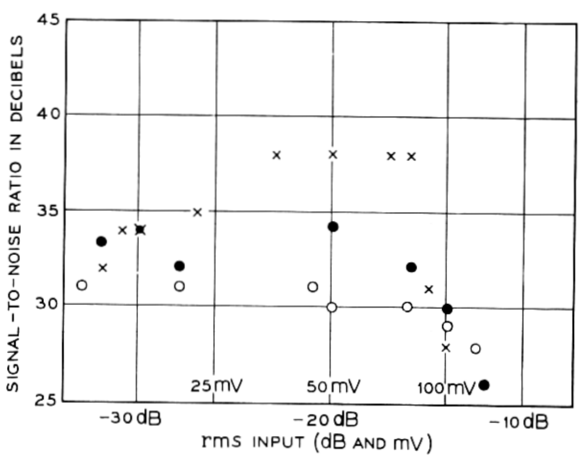

Fig. 11 — Subjective signal-to-noise ratios using a test chart.

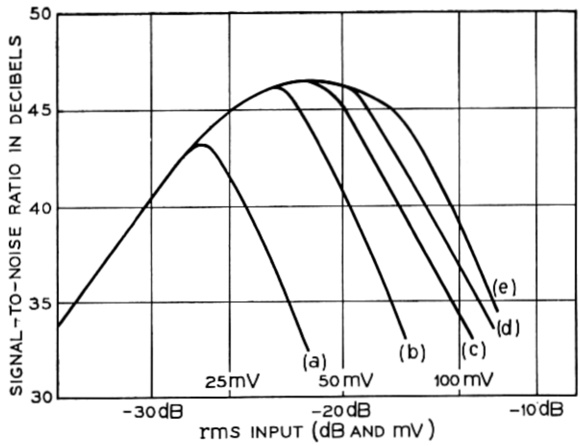

Fig. 12 — Subjective signal-to-noise ratios for various weight sequences by one observer: (a) unweighted; (b) 1, 1, 2 . . . 2; (c) 1, 1, 2, 3 . . . 3; (d) 1, 1, 2, 3, 4 . . . 4; (e) 1, 1, 2, 3, 5 . . . 5.

realistic evaluation of the coder than any objective measure we have used. Objective measurement, however, is needed for commercial evaluations. A useful method is a noise loading test.

4.3 *Noise Loading Tests*

Because of difficulty in characterizing video signals and human observers, some theoreticians have considered gaussian noise as the input when determining a coder's signal-to-noise ratio. Their results can be tested with a noise loading measurement. Such a measurement is described here in order to provide a comparison with published figures for other coders and to provide data for theoretical confirmation.

For these tests, gaussian noise with the spectrum shown in Fig. 7a was the coder input; the resultant output power was measured in selected 1 kHz bands. This power comprises a representation of the input with additional noise generated in the coder itself. A band rejection filter was then inserted before the coder to block the applied noise in the frequency band where the measurement is made; the measured power is therefore the noise generated in the coder alone. A signal-to-noise ratio for the coder can be determined from these two measurements. It is an objective measurement of the coder's properties in the particular band of frequency chosen.

Figure 15 gives the objective signal-to-noise ratio at 14 kHz for various weighting. These curves show that the weights have little ad-

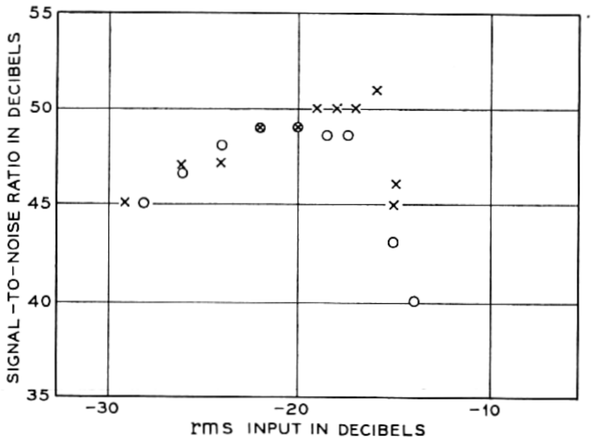

Fig. 13 — Subjective evaluation of two weight sequences: 0 = 1, 1, 2, 3, 5 . . . 5 and x = 1, 1, 2, 3·6, 4·7, . . . 4.7.

vantage for coding this noise. This is not surprising because the weights were chosen to suit the characteristics of video signals—especially the property that large values of the signal derivative occur as a few sharp spikes separated by relatively constant levels, whereas the derivative of the noise has a gaussian distribution. Figure

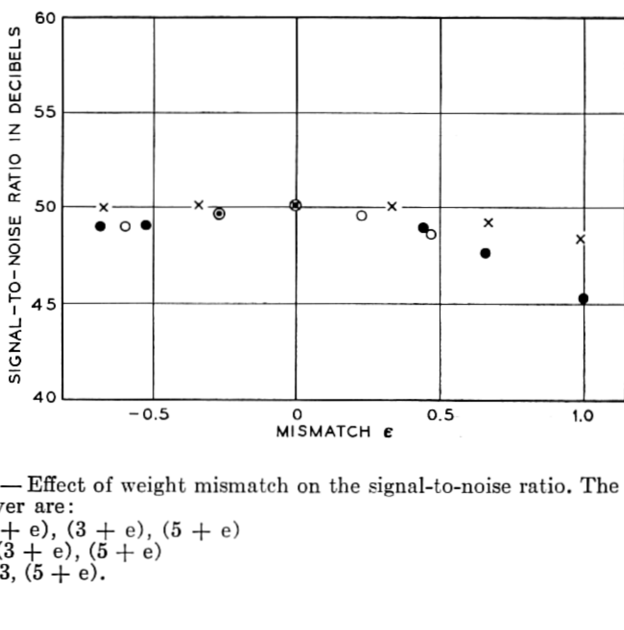

Fig. 14 — Effect of weight mismatch on the signal-to-noise ratio. The weights at the receiver are:
. 1, 1, (2 + e), (3 + e), (5 + e)
x 1, 1, 2, (3 + e), (5 + e)
o 1, 1, 2, 3, (5 + e).

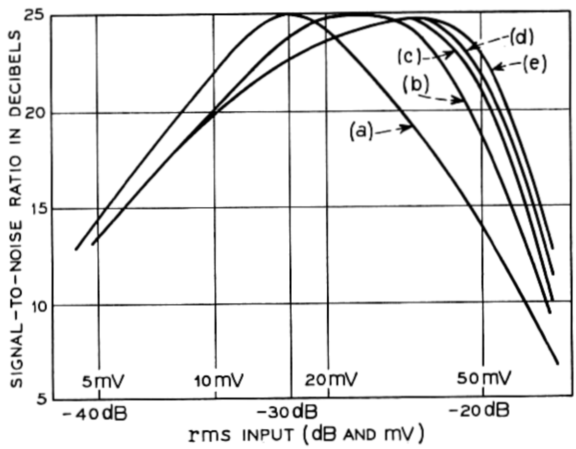

Fig. 15 — Noise loading results at 14 kHz. The weights are: (a) unweighted; (b) 1, 1, 2 . . . 2; (c) 1, 1, 2, 3 . . . 3; (d) 1, 1, 2, 3, 4 . . . 4; (e) 1, 1, 2, 3, 5 . . . 5.

16 shows how the signal-to-noise ratio depends on the frequency at which the measurement is made. By combining this result with the known spectrum of the input, it can be shown that the net signal-to-noise is about 22 dB.

Figure 17 shows signal-to-noise ratios obtained in the same way as those in Fig. 15, but using a video signal as input. These curves more

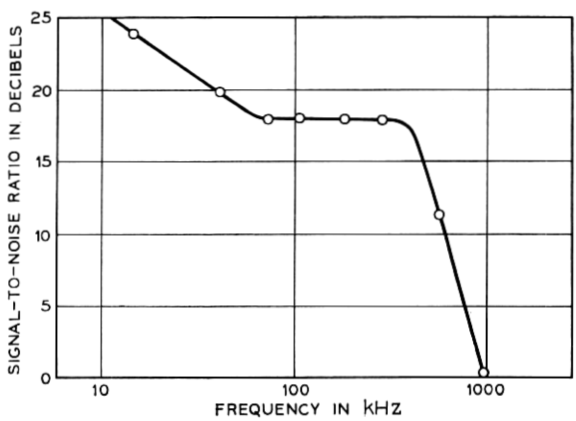

Fig. 16 — Loaded signal-to-noise ratios in 1 kHz slots at various center frequencies.

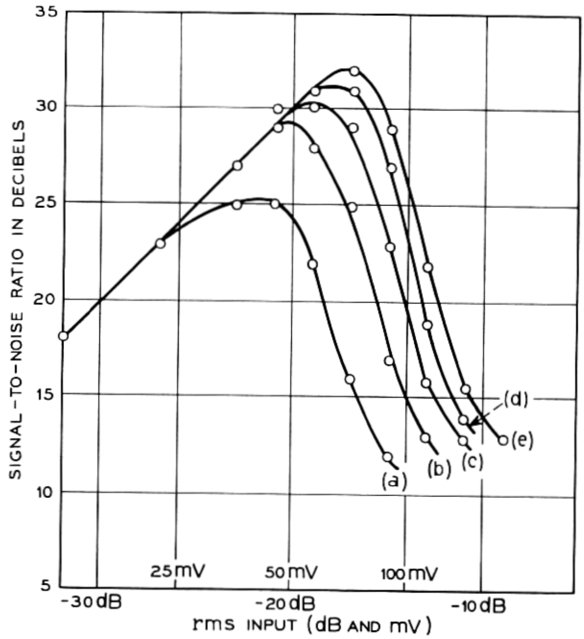

Fig. 17 — Objective signal-to-noise ratios at 14 kHz with video input, using the same weights as in Fig. 13.

nearly resemble the subjective results in Fig. 12. Notice that Fig. 12 refers to peak signals and Fig. 17 to root mean square signals; this accounts for the 11dB difference in the ordinates.

V. CONCLUSIONS

Weighting the step size of a one-bit coder improves the quality of the transmitted signal and broadens its tolerance to changes of input amplitude. The weighting is easily implemented with integrated circuits; in fact, the whole codec need be more complex or expensive than a simple radio receiver. The circuit tolerates up to ±30 percent mismatching of the transmitter and receiver (that is, $\epsilon = 0.3$ in Fig. 14). This is an important property for network applications where each transmission is available to many receivers.

The coder has been presented as a useful circuit for a particular application. No theoretical method for optimizing the companding is known because of difficulty in analyzing a system that incorporates an interaction of a television source, a human observer, quantization, linear filters, and digital processing. Instead, the circuit has been considered intuitively as an extension of the direct feedback coder described in Ref. 9. Indeed, the filters used are those recommended in that work.

We emphasize the tolerance of the circuit to parameter changes because attempts to improve a coder sometimes peak its response about certain parameter values. These parameters are then critical factors in the design. The present coder is very tolerant to changes; this is an important practical advantage.

When each element of a television signal is coded with three bits, the degradation of the picture is subjectively equivalent to about −50 dB of added noise. When the coder was adapted for voice transmission, telephone quality speech could be transmitted using a 50 KHz digit rate. In both examples the coders accepted a wide range (10 dB) of input level.

APPENDIX

The Circuit and Effect of Mismatched Weights

A.1. *The Circuit*

A.1.1. *Circuit Outline*

It is important that the transmission delay around the feedback loop not exceed a sample interval. Otherwise the excess delay will cause a low frequency instability called double moding. Correct operation requires that each decision of the threshold be sent around the feedback in time to fully influence the next decision. Meeting this requirement at high sampling rates is difficult but simplified by moving the weighting circuit, in Fig. 4, outside the feedback loop, as in Fig. 18. Now, each threshold decision activates a switch, *S,* that sends either of two values to the integrator. These two values have been set up by previous code values held in registers. For this purpose the threshold decision is placed in a flip-flop, *F*, in readiness for ensuing decisions.

A.1.2 *Circuit Action*

All the components of the feedback loop are dc coupled, enabling the levels in the circuit to be well defined and avoid displacements caused by spurious charges on coupling capacitors.

The timing cycle is given in Fig. 18. When gate T_1 conducts, it

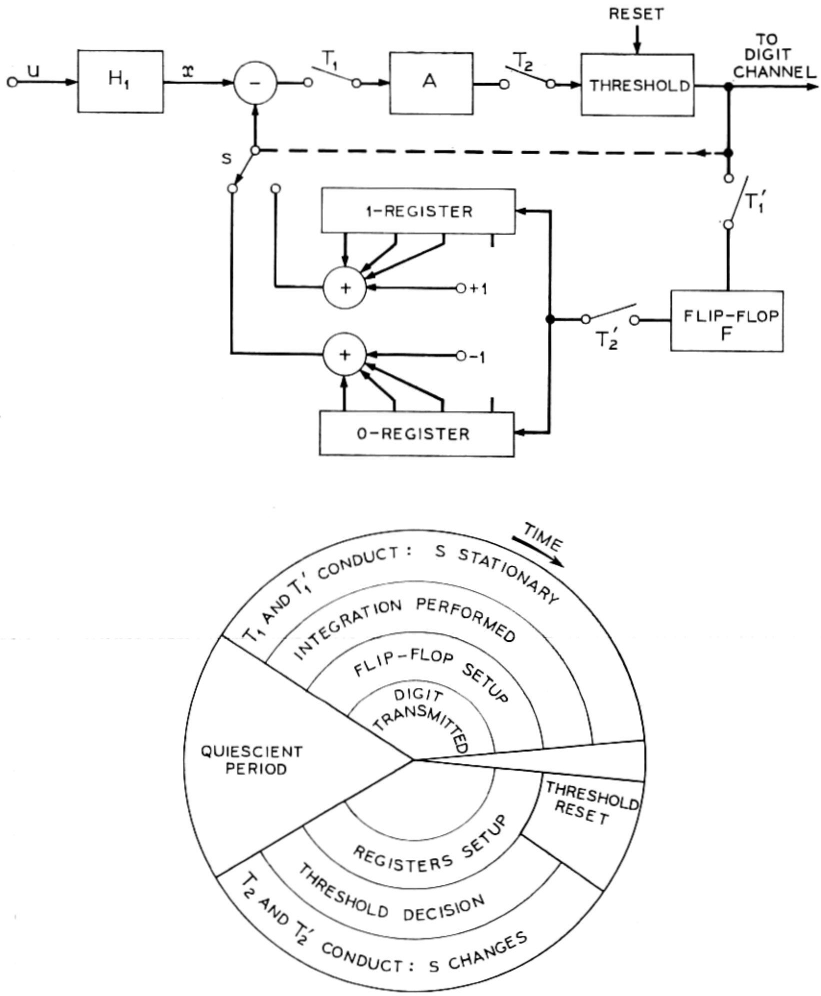

Fig. 18 — The block diagram and timing cycle.

samples the difference between the input and the feedback. It thus defines the pulse width fed to the integrator and isolates the integrator from the input while a polarity decision is being made on its output.

The second gate, T_2, conducts a short while after T_1 switches off. It defines the time in which decisions are made. The threshold circuit is bistable and so holds its decision until reset.[10] Resetting occurs just as T_2 starts conducting. A negative signal applied to the threshold input leaves it in the "off" state; a positive signal switches it on. Once on, the circuit cannot be switched off again from the input terminal.

The output from the threshold circuit sets the switch, S, in readiness for the next conduction of gate T_1. When T_1 conducts, the digit gate T'_1 also conducts, placing the decision in flip-flop F. At this time digit gate T'_2 is off, it conducts at the same time as T_2 transferring the content of F to the registers. A "one" in F resets the 0-register and inserts a "one" into the one-register, shifting up its content. Similarly, a "zero" in F resets the one-register and inserts a "one" into the 0-register. These registers feed signals to adding circuits whose outputs provide the quantized signals, either of which is selected by the next decision, using switch S.

A.2 *Effect of Mismatched Weights*

Any codec needs a digital-to-analog converter at its receiver to assign analog values to the digital code. For the ordinary one-bit codec it is simply a pulse shaping circuit; for multilevel codecs it is more complex, because a variety of different analog values must be generated in response to different code words. The present codec uses a digital-to-analog converter with eight outputs, ± 1, ± 2, ± 3, ± 5 corresponding to different code patterns.

What happens when there is an error in one of the levels generated at the receiver? Every time the code calls for that level, the output will be wrong. When use of each level is completely determined by the instantaneous input the error is a distortion, or nonlinearity, of the output. This is a characteristic of straight pulse code modulation. Conversely, when use of a particular level is not determined by values of the input, but is used almost at random, then errors in it appear as noise on the output. This often happens in multilevel differential and feedback coders. For the companded one-bit coder described, there appears to be a high correlation between amplitudes of the pre-emphasized input x and use of particular levels. Mismatching weights are thus, approximately equivalent to a distortion of x.

Distortion of the pre-emphasized signal appears on the output as a distortion of edges and busy areas. The errors persist for about 3 microseconds which is the time constant of the de-emphasizing filter. The visible effect of small errors is not displeasing; it resembles a change of contrast in the busy areas, and sometimes, a little streaking near the edges.

If the weight sequence used is one that makes the coder unstable, then the correlation between values of x and use of particular weights is lost, and mismatching weights introduces noise.

VI. ACKNOWLEDGMENT

We gratefully acknowledge helpful discussions with our colleagues in Departments 135, 321, and 462 at Bell Telephone Laboratories.

REFERENCES

1. O'Neal, J. B., "Delta Modulation Quantizing Noise, Analytical and Computer Simulation Results for Gaussian and Television Input Signals," B.S.T.J., *45,* No. 1 (January 1966), pp. 117–141.
2. O'Neal, J. B., "Predictive Quantization Systems for the Transmission of Television Signals," B.S.T.J., *45,* No. 5 (May-June 1966), pp. 689–721.
3. O'Neal, J. B., "A Bound on Signal-to-Quantizing Noise Ratio for Digital Encoding Systems," Proc. IEEE, *55,* No. 3 (March 1967), pp. 287–292.
4. Greekes, J. A., and de Jager, F., "Continuous Delta Modulation," Philips Research Report 23, 1968, pp. 233–246.
5. Tomozawa, A., and Kaneko, H., "Companded Delta Modulation for Telephone Transmission," IEEE Trans. on Communication Theory, *16,* No. 2 (February 1968), pp. 149–156.
6. Brolin, S. J., and Brown, J. M., "Companded Delta Modulation for Telephony, IEEE Trans. on Commun. Theory, *16,* No. 2 (February 1968), pp. 157–162.
7. Abate, J. E., "Linear and Adaptive Delta Modulation," Proc. IEEE, *55,* No. 3 (March 1967), pp. 298–307.
8. Winkler, M. R., "Pictorial Transmission with H.I.D.M.," IEEE Int. Conf. Record, Part 1, 1965. pp. 285–290.
9. Brainard, R. C. and Candy, C. J., "Direct Feedback Coders: Design and Performance with Television Signals," Proc. IEEE, *57,* No. 5 (May 1969), pp. 776–786.
10. Candy, C. J., unpublished work.