B. S. T. J. BRIEF

Interpolation of Data With Continuous Speech Signals

By M. R. SCHROEDER and S. L. HANAUER

(Manuscript received March 1, 1967)

In some communications systems, the need arises for temporally interpolating data or signalling information during continuous speech.[1] If the required time gaps are created by simply interrupting the speech signal, severe degradation of speech quality and some loss in intelligibility results.

The reason for the degradation is twofold:

(*i*) The interruptions introduce *discontinuities* in the speech signal—two for every interruption.

(*ii*) The interruptions, unless occurring pitch synchronously, create an *inharmonic* signal.

In the following, a proposal is described which avoids discontinuities and is pitch synchronous—without the need for pitch detection. Average "off-time" ratios of 30 percent have been achieved for *continuous* speech without audible degradation. These results were obtained by computer simulation of a sampled data system. The instrumentation for a real-time analog system is simple.

The gaps created by this method occur at irregular intervals in time. Thus, for a steady flow of data or signalling information, some buffer storage and coding that distinguishes "gaps" (interpolated data) from speech is required.

In the proposed interpolation system, the speech signal $s(t)$ is divided by its envelope

$$a(t) = [s^2(t) + \hat{s}^2(t)]^{\frac{1}{2}}, \tag{1}$$

where $\hat{s}(t)$ is the Hilbert transform[2] of $s(t)$. The resulting signal is then multiplied by a modified envelope

$$\tilde{a}(t) = \{a(t) - c\}_+ \equiv \begin{cases} a(t) - c & \text{if} \quad a > c \\ 0 & \text{if} \quad a \leqq c, \end{cases} \tag{2}$$

where the function $\{\ \}_+$ equals its argument for positive arguments and is zero otherwise.

The combination of these two operations results in the desired interrupted signal

$$s_i(t) = \frac{s(t)}{a(t)} \{a(t) - c\}_+ . \tag{3}$$

The average off-time ratio depends on the magnitude of the constant c. For Gaussian signals, this ratio is given by

$$r_{\text{off}} = 1 - \exp\left[-c^2/\overline{a^2}\right]. \tag{4}$$

In one of the computer simulations, c was chosen equal to 0.5 $\bar{a}$. For a Gaussian signal, this choice corresponds to

$$c = \tfrac{1}{4}[\pi \overline{a^2}]^{\frac{1}{2}}. \tag{5}$$

Thus, the average off-time becomes

$$r_{\text{off}} = 1 - \exp\left[-\pi/16\right] = 0.18 = 18\%. \tag{6}$$

The actually observed off-time ratios for $c = 0.5\,\bar{a}$ for two test sentences were 26 percent for male speech and 16 percent for female speech. ($\bar{a}$ was obtained by averaging $a(t)$ over 20 msec with a rectangular time window.)

Fig. 1 shows microfilm outputs from a computer simulation. The constant c was chosen equal to 0.9 $\bar{a}$ in order to achieve off-time ratios near 50 percent. The first line shows the original signal $s(t)$, the second line the interrupted signal $s_i(t)$ and the third line the "switching function" $\{1 - c/a(t)\}_+$. The actual off-time ratio for the total utterance (one sentence, male speaker) was $r_{\text{off}} = 55$ percent.

The speech quality for this rather large off-time ratio was judged somewhat nasal but as intelligible as the original. It is possible that

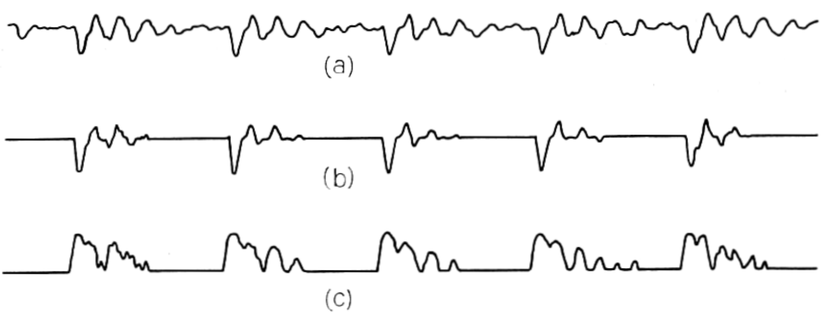

Fig. 1 — (a) Original signal $s(t)$. (b) Interrupted signal $s_i(t)$. (c) "Switching function" $\{1 - 0.9\bar{a}(t)/a(t)\}_+$

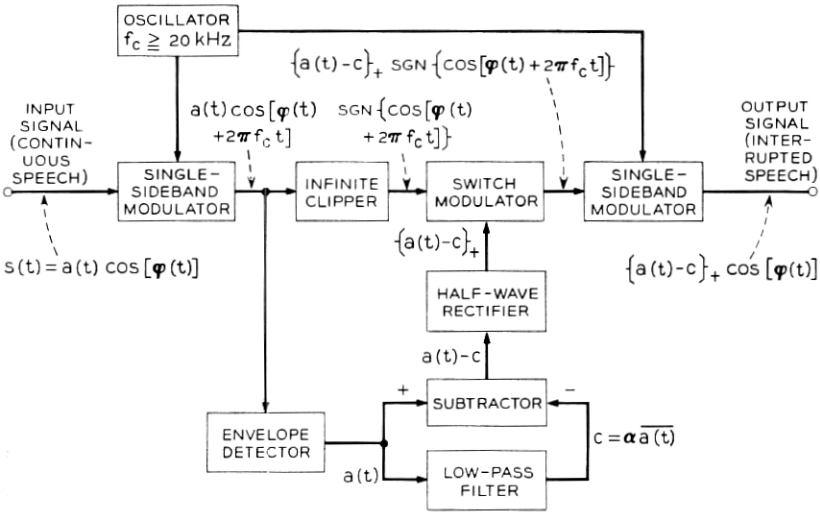

Fig. 2 — Possible implementation of a system for interpolation of data with continuous speech signals.

even better results might be achieved by "smoothing" the switching function, by filling in short gaps and by eliminating short speech bursts.

A possible implementation is shown in Fig. 2 in block diagram form. The division by $a(t)$ indicated in (3) can be effected by single-sideband modulation followed by infinite clipping. The function $\{\cdot\}_+$ corresponds to standard half-wave rectification. The multiplication by $\{a(t) - c\}_+$ can be effected by a switch-type modulator. The desired interrupted signal is obtained by a downward frequency shift of the modified-envelope single-sideband signal.

The problem of interpolating signalling information with speech arose in a new mobile communication system for trains and was brought to our attention by Mr. C. E. Paul.

REFERENCE

1. Paul, C. E., Simultaneous Noninterference Transmission of Continuous Speech and Data, unpublished work.
2. Kaplan, Wilfred, *Operational Methods for Linear Systems,* Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962, pp. 395–400.