A tone relay can eliminate troublesome signaling tone distortion and other problems in packet data networks that employ speech compression.

Passing Tones in Voice-over-Packet Data Systems

By Scott Kurtz

Low-rate speech compression algorithms for voice-over-packet data networks can often distort signaling tones excessively. Tone relays and tone relay algorithms especially developed for voice-over-packet and other systems that employ speech compression enable you to avoid the problem. These algorithms address a number of subtle issues that affect robustness and performance. (Even when using one, however, you need to do a significant amount of testing under a variety of conditions to ensure robustness.) A good one should be flexible enough to handle the right set of signaling tones yet simple enough to integrate into a host application.

The idea of carrying speech over packet data networks is gaining acceptance in the world of telephony. Many standards bodies have scrambled to determine the best way to do that over the various types of packet data networks, like IP, ATM, and frame relay. All packet networks have limited bandwidth; for that reason, speech compression is an essential ingredient of voice-over-packet standards, in order to make the best use of the channels.

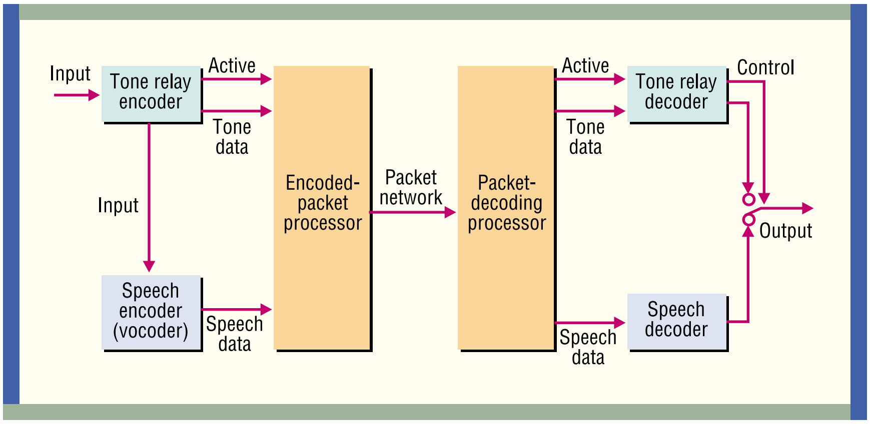

Figure 1. A communications system can employ a tone relay working in parallel with its speech coder. When tones are detected, the tone information is sent over the same communications network as the encoded speech data, sometimes replacing the speech data. At the decoding end, the tone relay decoder regenerates the tones, and the output replaces the decoded speech.

Speech compression algorithms remove redundancy from speech data by extracting key information from speech signals. This information often includes parameters that model the human vocal tract. A good model requires only a few bits to specify the parameters while still providing good speech quality when the speech signal is regenerated. In general, a higher degree of compression (lower bit rate) results in lower speech quality.

Unlike Morse and Huffman coding, speech compression algorithms result in the loss of information. Although good algorithms hold down the loss in perceived speech quality, the lower-rate algorithms aren't adept at passing many non-speech signals, including DTMF, MF R1, MF R2 Forward and Reverse, and Call Progress tones. In fact, many lower-rate speech compression algorithms distort signaling tones beyond the point of reliable detection.

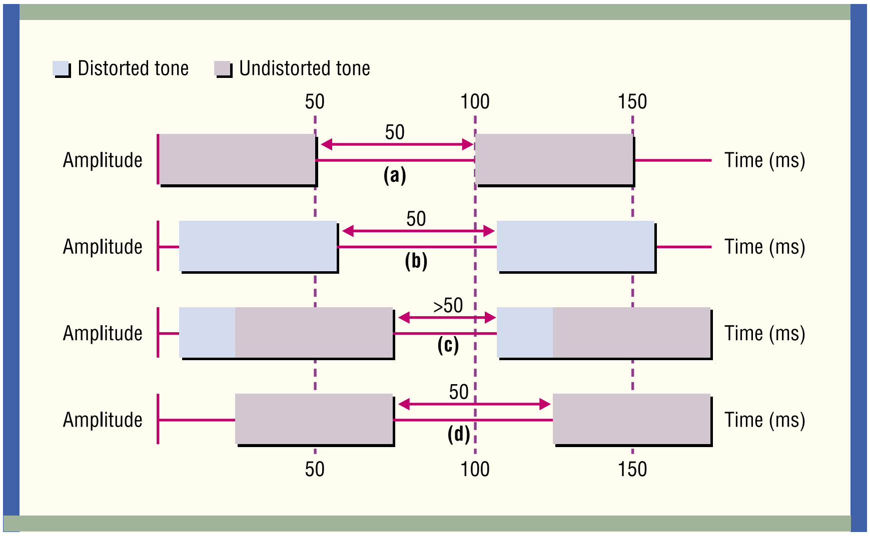

Figure 2. ***Too long a segment of tone passing through the vocoder can lead to adverse effects. Here, for an input sequence of two 50-ms tone pulses (a), the vocoder delays and distorts the signal (b). The duration of the pulses has increased in the decoded output, reducing the interdigit time, and the initial part of the pulses is still distorted (c). If the interdigit time is not met, the detector at the far end can make errors. Using leading-edge suppression eliminates the distortion and maintains the interdigit time (d).***

SYSTEM OVERVIEW

A tone relay implementation consists of two layers of functionality: the relay encoding and decoding and the underlying tone detection and generation, as shown in Figure 1. The tone relay has a component that operates at the encoding side of the link, the tone relay encoder, and one that operates at the decoding side, the tone relay decoder. In the figure, the input is a sequence of a sample representing the voiceband signal. Normally, the sampling rate is 8,000 samples per second.

The input samples feed both the speech encoder and the tone relay encoder. The speech encoder compresses the speech data, and the tone relay encoder detects the presence or absence of a signaling tone. If a tone is present, the tone relay encoder encodes information about the tone that enables it to be reproduced at the other end of the communications link. The tone relay encoder also issues an Active flag that, when set, indicates that valid tone data is present.

The speech data, tone data, and Active flag form the inputs to the encoded-packet processor, which creates a speech or tone packet in accordance with the appropriate voice-over-packet specification. The packet then goes to the opposite end of the link via the packet network.

The packet-decoding processor receives packets from the packet network and sends compressed speech data to the speech decoder. In some systems, speech data isn't sent when tone data is sent. If that's the case, the processor informs the speech decoder that the frame is missing and the speech decoder acts accordingly. The processor also sends the tone data and the Active flag to the tone relay decoder. If tone activity is present, the tone relay decoder regenerates the original tone based on the parameters included in the packet.

A switch determines whether the decoded speech or the regenerated tone data serves as the output. The tone relay decoder controls the switch, which is set to the speech decoder unless a tone is being regenerated.

TONE PASSER CHARACTERISTICS

On the surface, the design of a tone passer might appear straightforward. There are, however, some implementation details, such as frame size, that can have a significant effect on its performance.

Speech compression algorithms operate on a frame-by-frame basis. The input speech is therefore divided into frames. Each frame contains a given number of samples, as defined by the algorithm. Typical frame lengths and their corresponding sample counts are 2.5 ms (20 samples), 10 ms (80 samples), and 30 ms (240 samples).

To reconstruct signaling-tone

bursts more precisely, a tone relay should use finely specified durations. Consequently, the tone relay encoder must be able to analyze the input signal using small frame sizes so that the burst length is quantized in sufficiently small intervals.

Although a smaller interval allows a more accurately reconstructed tone, more processing power is needed to detect the smaller intervals. A trade-off is therefore necessary.

A good tone relay should generate tone bursts with an accuracy of ±3 ms. The tone burst interval may not coincide with the frame size of the vocoder or with the analysis frame size of the tone detector in the tone relay encoder. Accordingly, the tone relay algorithm must work out those differences.

Amplitude quantization is another important detail. Note that the tone relay quantizes both the time interval and the signal amplitude. Tone detectors in the network must detect signals over a wide range of amplitudes. Nevertheless, a tone relay should be designed so that the amplitude of the regenerated signal remains close to that of the original signal. The telephone network already contains losses, and the tone relay shouldn't contribute any more. If a network is already near the upper limit of signal loss, the tone relay could push it over the top if the regenerated amplitude isn't sufficiently accurate.

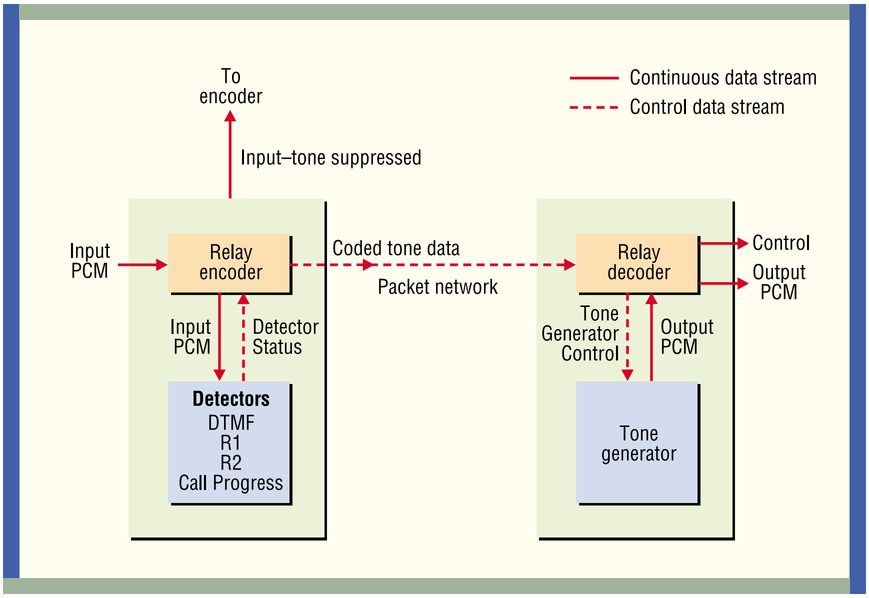

Figure 3. The tone relay encoder function includes tone detectors, which supply status information the encoder uses to produce coded tone data. Similarly, the tone relay decoder function includes a tone generator. The tone generator uses control information from the relay decoder to re-create the tone.

A good tone relay should be able to quantize the signal amplification with quantization intervals of 3 dB or less.

Once again, there's a trade-off. As the number of amplitude quantization levels increases, the bandwidth required to pass the quantized levels increases. A good tone relay should be able to quantize the signal amplitude with quantization intervals of 3 dB or less.

The dynamic range of the amplitude quantizer should span the range called for by the associated signal detector specification. For example, if the detector specification indicates that a detector must detect signals between 0 and –25 dBm, the tone relay amplitude quantizer should span that range.

Leading-edge suppression is another parameter to worry about. To detect a signaling tone reliably without excessive probability of a false alarm, many samples of input data must be analyzed before deciding whether a signaling tone is present. The tone relay encoder output is therefore delayed with respect to the onset of the signaling tone. In a system that employs tone relay, the leading edge of the signaling tone could possibly pass through the vocoder before the tone relay detects the tone.

If a long enough segment of tone passes through the vocoder, two adverse effects may result, as shown in Figure 2. Here, a sequence of two 50-ms tone pulses are the input to the encoding side of a communication link (Figure 2a). The vocoder introduces both delay and distortion. (Figure 2b).

The first effect is that the tone burst that emerges from the decoding side is extended in duration, and distortion remains at the start of the tone bursts (Figure 2c). The beginning of the output signal is the portion passed by the vocoder, and the tone relay decoder generates the remainder. As a result, the interdigit time is shortened when signaling tones occur in rapid succession. If the interdigit time isn't met, the detector at the far end can make errors. Leading-edge suppression eliminates the effect, as well as the distortion (Figure 2d).

The second effect is similar to the first. Since it's presumed that the vocoder distorts the signaling tone, the portion that it passes is distorted, but the portion that the tone relay passes is not. However, a phase discontinuity occurs between the two portions of the output signal. The discontinuity could cause the detector at the far end to detect a split digit or two separate digits for a single input digit.

Consequently, a tone relay should be capable of detecting the leading edge of a signaling tone as soon as possible. When the leading edge is detected, a tone suppressor should be used to remove the signaling tone's frequency components from the signal going to the vocoder.

Tone suppression could be done more crudely by muting the input to the speech encoder. However, although muting is a simpler solution, it isn't desirable. The leading-edge detector isn't as robust as the detector itself because it must make a decision on a shorter-duration signal. The leading-edge detector is therefore more prone to false alarms. Muting the input to the speech encoder when the leading-edge detector goes off makes it more likely that speech will be muted, rather than the signaling tone. So it's better to suppress only the frequency components that are part of the detected signaling tone. That way, if speech is present, the speech signal isn't changed too much.

FROM THEORY TO PRACTICE

Now let's take a closer look at a tone relay implementation. Figure 3 breaks down the architecture shown in Figure 1.

The input PCM signal is fed into the tone relay encoder, which in turn feeds it to the various types of tone detectors. The tone detectors return status information including tone presence, early tone detection status, and signal amplitude. The tone relay encoder monitors the detector status information and sends coded tone data to the communications channel. The coded tone data includes information indicating which tone (if any) is present, the level of the tone, and the time elapsed since the start of the tone. The relay encoder also produces a suppressed version of the input PCM signal for use by the speech encoder.

The coded tone data is fed to the relay decoder. The decoder uses the data to control the tone generator with a Tone Generator Control signal, which contains frequency, amplitude, and pulse duration information. The tone generator synthesizes the tone and sends it back to the relay decoder.

A tone relay should be capable of detecting the leading edge of a tone as soon as possible.

A carrier-class DSP-based tone relay software package can demonstrate how tone relay functionality can be integrated into a host application. For the discussion of such software, go to http://edtn.com/cs/EE/tonerelaysw.fhtml ◆

Scott Kurtz (skurtz@adt-inc.com) is the vice president of engineering at Adaptive Digital Technologies, Inc. in Conshohocken, Pa., where he has focused on the development of the company's DSP based products and technology. He has specialized in digital signal processing and digital communications over the course of his 18-year career, beginning at RCA Corporation's Government Communications Systems Division. At InterDigital Communications Corporation, he was instrumental in developing the first digital wireless local loop telephone system, which was the precursor to today's digital cellular telephone systems. He holds six patents, with two more pending.