N88-25755

DEVELOPMENT OF AN 8000 BPS VOICE CODEC FOR AvSat

JOSEPH F. CLARK, AvSat Program, Aeronautical Radio, Inc., United States

AERONAUTICAL RADIO, INC.
2551 Riva Road
Annapolis, Maryland 21401
United States

ABSTRACT

Air-mobile speech communication applications share robustness and noise immunity requirements with other mobile applications. The quality requirements are stringent, especially in the cockpit where air safety is involved. Based on these considerations, a decision was made to test an intermediate data rate such as 8.0 and 9.6 kb/s as proven technologies.

A number of vocoders and codec technologies were investigated at rates ranging from 2.4 kb/s up to and including 9.6 kb/s. The "proven" vocoders operating at 2.4 and 4.8 kb/s lacked the noise immunity or the robustness to operate reliably in a cabin noise environment. One very attractive alternative approach was Spectrally Encoded Residual Excited LPC (SE-RELP) which is used in a multi-rate voice processor (MRP) developed at the Naval Research Lab (NRL). The MRP uses SE-RELP at rates of 9.6 and 16 kb/s. The 9.6 kb/s rate can be lowered to 8.0 kb/s without loss of information by modifying the frame. An 8.0 kb/s vocoder was developed using SE-RELP as a demonstrator and testbed. This demonstrator is implemented in real-time using two Compaq II portable computers, each equipped with an ARIEL DSP-16 Data Acquisition Processor.

INTRODUCTION

The Airlines Electronic Engineering Committee (AEEC) has undertaken development of a "characteristic" for satellite communications equipment — in effect a guideline for equipment and airframe manufacturers, aviation users, and service providers. Designated Project 741, it is being developed in parallel with a wide variety of intense activities in the mobile satellite communications area.

The primary schedule constraint for the AEEC has been introduction of a new generation of aircraft to be delivered "satcom ready" starting in 1989. As a consequence, the AEEC formulated a three-phase schedule for Project 741. By late 1986, the Satellite Subcommittee of the AEEC formed a working group to recommend a voice coding standard. Although there has been a hiatus in the activity of that working group, its efforts both in fulfilling its mission and in assuring that all aspects of the voice communication system be considered together are continuing.

This paper reports on the specific effort to develop a voice coding standard using the transmission rate of 8.0 kilobits per second (kb/s) and focuses on the similar performance, as shown in testing, of the 8.0 kb/s and 9.6 kb/s codecs available today.

ADAPTING TO THE USER COMMUNITY

The aviation community is loosely divided between "general aviation" and commercial passenger aviation. The aircraft used vary in size from the very large to the very small. The length of flights varies from less than one hour to more than eight hours. Although data communications increasingly is the mechanism for reporting aircraft status and for routine operational exchanges, one common denominator is a requirement for cockpit voice — a communication path for the flight crew to talk with ground crews.

While the primary volume application of voice appears to be passenger telephony, the cockpit voice quality requirements may be the primary determinant of the overall requirements because of the ties to safety.

While only experimental equipment has been installed to date, it is clear that aviation will use both high-gain and low-gain antennas. The service providers will be using a mix of global beams and spot beams from the satellite. As a result, three classes of service can be provided: low-rate data; medium-rate data and low-rate voice; and the nominal, which is medium-rate data and "near-toll-quality" voice.

We suggest that a single algorithm for both low-rate and medium-rate voice would be the preferred solution. This approach was developed by the Naval Research Laboratory (NRL)[1] and is in operation today. The method for providing voice communication through a low-rate data service is to buffer short messages coded at the nominal voice rate, and transmit them as extended messages at the actual channel data rate.

PERFORMANCE REQUIREMENTS

So far, no one has defined well enough the measurable parameters for repeatable measurements of "near toll quality." However, the achievable performance of codecs is well known. Figure 1 depicts the achievable performance of a wide variety of codecs as determined in tests. But in order to specify a required performance, it is necessary to determine minimum acceptable performance in terms of parameters that can be measured in a consistent way. Unfortunately, voice quality is by its very nature subjective. It can only be quantified as a statistic derived from inherently variable data.

The performance requirements endorsed by the AEEC Voice Working Group[2] are simple:

- Subjective evaluation measured by Mean Opinion Score - midway between Fair and Good (3.25 on a scale of 1-5).
- Intelligibility measured by Rhyme Test - 87 or better.
- Conversational evaluation - 75 or better.
- Delay measured with codecs back to back - no more than 65 milliseconds.

- Performance measured with a channel bit error rate (BER) of 1 in 1000 bits.

In addition, the group recognized that the operational environment would include low-rate data transmissions, rather severe outages and interfaces with established terrestrial telephone networks. It was decided that data should be routed around the codecs. Not only does the requirement for coding data as well as voice put significant demands on the coding algorithm, it implies significant additional acceptance testing.

The new CCITT coding standard for digital transmission of wideband audio signals, G.722, provides a similar bypass mechanism for user data. The separation of voice and data allows a better system design[3].

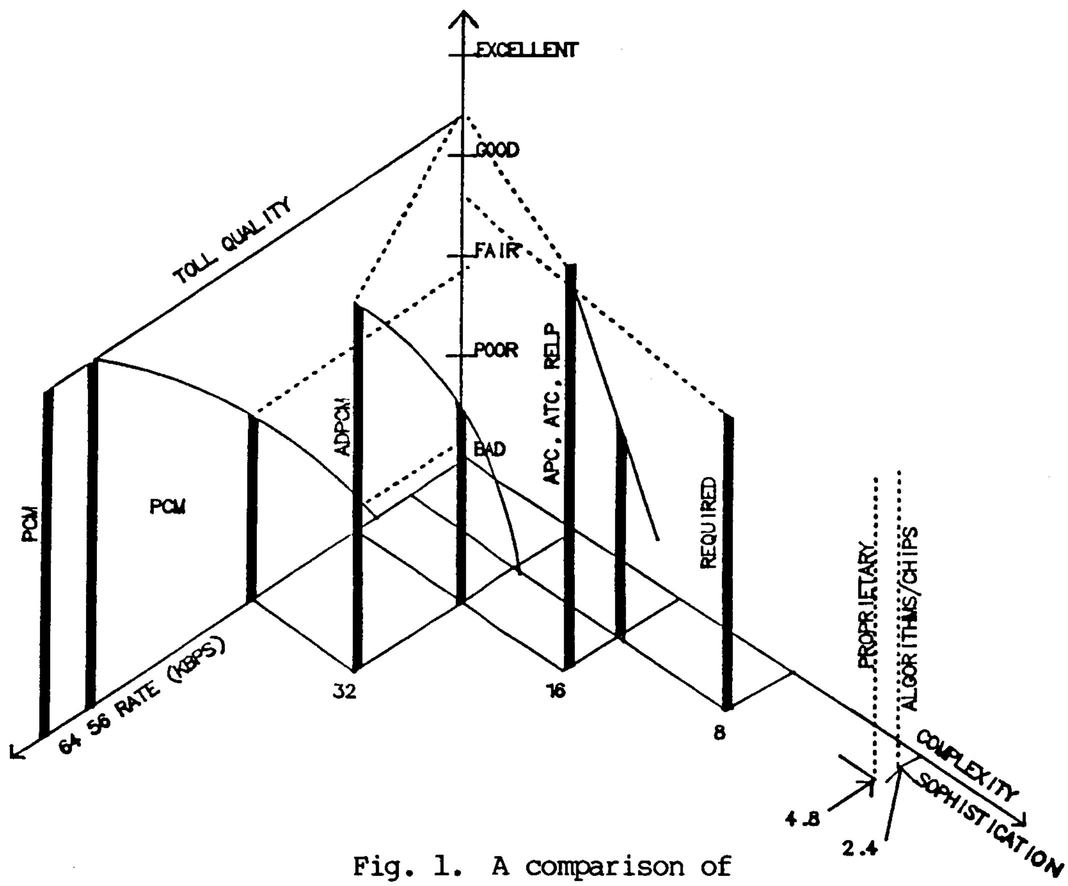

Fig. 1. A comparison of speech transmission

APPROACHES

A CCITT standard ADPCM would simplify the task, but there is more complexity in the transmission channel. Limitations of the medium, and power and bandwidth constraints drive the choice of rates to the lowest possible. On the other hand, the realizable quality at a given rate drives the choice upwards. The result then, is a compromise on an intermediate rate that approximates the desired performance without devouring all available capacity.

The rate issue has resolved to a choice between 8.0 kb/s and 9.6 kb/s. The 8.0 kb/s rate is based on compatibility with expected subchannelling of the ISDN basic 64 kb/s communication rate. Intuitively, there should not be much difference in performance at these rates relative to, say, performance at 16 kb/s or at 4.8 kb/s. This intuition is supported by the evaluations received so far. (See Figures 2 and 3.)

Furthermore, performance differences seem to be related more to implementation than to algorithm. This result is quite reasonable since most of the algorithms are variations on the idea of using residual excitation with a linear prediction filter.

The conclusion then is that the AEEC's voice working group could not recommend a specific algorithm, or even a data rate. Only implementations of voice systems can be evaluated and recommended.

From an overall systems point of view, there is a significant difference between these rates. Choosing 9.6 kb/s without gaining a significant performance advantage is wasteful. In the best of proposed implementations, 9.6 kb/s uses $1.20 worth of resources for every $1.00 used by 8.0 kb/s. For SCPC voice channels as proposed in Project Paper 741, the 9.6 kb/s implementation requires almost twice as much bandwidth as the 8.0 kb/s implementation. (See Table 1.)

Table 1. Voice Channel Frame Parameters*

Voice Rate (bits)	FEC Rate	Channel Rate (bit/s)	Channel Spacing KHz	Bits in Frame		
				Voice/25 (v)	Data/96 (SU) (n)	Dummy (d2)
9600**	0.50	21000	17.5	192	3	84
8000**	0.75	12600	10.0	160	6	8

*Extracted from Project Paper 741, Part 2.
**Preferred Voice Channels

The AEEC voice working group with the assistance of the Boeing Airplane Corporation developed a standard test tape. This tape was distributed to approximately twenty-five organizations which had expressed an interest either in developing codecs for the aviation industry or in evaluating voice codecs. The tapes were fed in and out of candidate codecs, and the resulting tapes analyzed in listening tests conducted by Dynastat.

ONE 8.0 KB/S SOLUTION

There were two major concerns in dealing with established manufacturers of voice codecs. First, a satisfactory licensing agreement for the algorithm and implementation must be available to all avionics manufacturers. Second, manufacturers might continue to focus on products designed only for their terrestrial markets, i.e., codecs designed

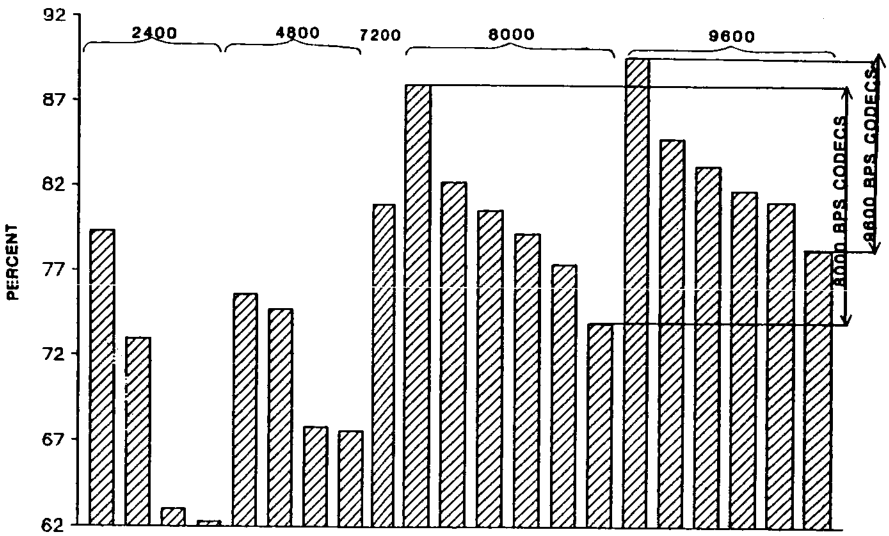

Fig. 2. Intelligibility: Cockpit
Noise Environment

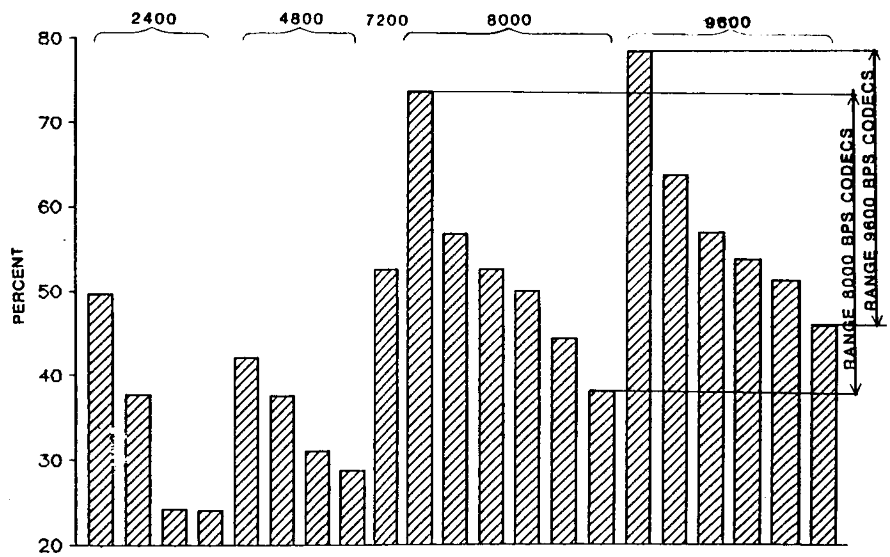

Fig. 3. Probability of User Acceptance Based
Based on Cockpit Noise Environment

for 2.4 kb/s, 4.8 kb/s, and 9.6 kb/s even though 8.0 kb/s is a demonstrably better choice for the emerging aviation industry characteristic.

In view of these concerns, ARINC contracted with Techno-Sciences, Incorporated (TSI) of Greenbelt, Maryland, to develop an 8.0 kb/s voice codec from information published in the public domain. The premise was: if a credible voice codec could be developed by one or two key people working with limited time and budget, then avionics manufacturers could derive non-proprietary sources for voice codecs, and on a realistic schedule with confidence.

TSI, with ARINC assistance, developed a pair of 8 kb/s voice codecs operating in real time. The codecs are connected through the serial (RS232) ports of two Compaq Portable II computers. Speech is sampled 8000 times a second and blocked into groups of 180 samples. Ten reflection and filter coefficients are calculated using autocorrelation. The quantized reflection coefficients are used to form the filter which then is used to generate the prediction residual.

A 96-component spectral representation of the prediction residual is computed using the Winograd algorithm. These components are quantized and coded into a 180-bit block with the reflection coefficients. A new block is generated and transmitted 44.4444 times per second. The key factors in the codec performance are the reflection coefficient calculation and the number of transmitted spectral components. The current codec transmits 126 spectral components, the same number transmitted in the NRL 9.6 kb/s rate codec.

After 6 months of development, a codec emerged that is credible but admittedly is incomplete and needs improvement. Several areas of the basic LPC algorithms have been tagged for further work in a second phase of the project. In the second phase of the codec development, vector coding of the spectral components will be introduced. There is good confidence in the expected results.

CONCLUSION

The AEEC Voice Working Group, with cooperation of the Boeing Airplane Company, the Federal Aviation Administration, the Rome Air Development Center and others, has developed a set of standards for voice codecs in the air mobile environment, collected a representative set of samples for a variety of codecs using a uniform test tape, and received an unbiased formal evaluation of the sample tapes. The results of the evaluation indicate that at least 7.2 kb/s speech coding will be required to reliably meet requirements in the aviation satellite environment. These results (see Figures 2 and 3) also indicate that performance at 8.0 kb/s is in the same range as performance at 9.6 kb/s — making 8.0 kb/s the better overall choice for a power and bandwidth-limited satellite environment to be globally interconnected with ISO-conforming networks.

REFERENCES

1. Kang, G. S. and Fransen, L. J., Second Report of the Multirate Processor for Digital Voice Communications, NRL Report 8614, September 30, 1982.
2. AEEC Letter 87-014/SAT-33, p. 11, Aeronautical Radio, Inc., 2551 Riva Road, Annapolis, MD 21401, February 13, 1987.
3. Mermelstein, Paul, G.722, a New CCITT Coding Standard for Digital Transmission of Wideband Audio Signals, <u>IEEE Communications</u> Magazine, Vol. 26, No.1, January 1988.
4. AEEC 700 Quick Check #50, November 4, 1987.
5. Techno-Sciences, Inc., Report No. 870909, 8 kbps Speech Demo Project, prepared for Aeronautical Radio, Inc., September 9, 1987.