Lead Feature

Digital Voice Communication

By Andy Talbot, G4JNT *

THIS EXPERIMENT started after a phone conversation with Charles Brain, G4GUO, where we had been discussing various data modes used by the military and other commercial users on the HF bands. One of the modes that has been in use for several years is digitised voice, transmitted in a bandwidth comparable with normal analogue voice communications and making use of existing transmitters and receivers. I must have said something like "it would be nice to try this on the amateur bands", because after the phone call finished Charles then went away and had a long think.

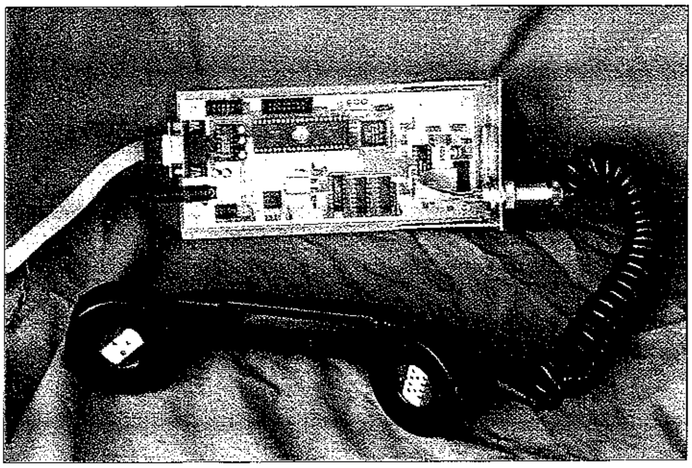
DIGITAL TECHNIQUES

MEANWHILE, WE need to digress a little to cover digital communication techniques and the problems given by the HF communications environment in order to fully appreciate why one type of datacomms technique is employed over any other.

When properly implemented, digital communications can show considerable advantages over their analogue counterparts with regard to quality and robustness over noisy transmission media - compare the quality of CD music recordings with the old vinyl or tape system and the new digital telephone network with the old system. But there are several major issues to be resolved before the conversion is made.

To digitise an analogue signal such as voice, it first has to be sampled, ie turned into a series of numerical values. Theory dictates that the sampling has to be at a rate at least twice that of the highest frequency component present - the Nyquist Criteria - and there must be no components present at more than half the sampling rate otherwise these will appear as spurious components at other frequencies causing distortion. This is known as aliassing, and the high frequency components need to be removed by conventional filtering before digitisation. For a voice signal as transmitted using telephone or SSB, the frequency range of 300-3300Hz is usually taken and dictates a sampling rate of at least 6.6kHz. In practice, to ease the anti-aliassing filtering, a sampling rate of 8000 is often adopted.

Since the analogue signal can have an infinite number of instantaneous amplitude levels, these cannot be represented exactly and it is necessary to choose a suitable number of levels to be used to represent the signal (instead of levels, it is more convenient to think of the number of binary bits needed to give the quantising needed, eg 8 bits gives $2^8 = 256$ absolute levels, 16 bits per sample gives $2^{16} = 65536$ levels). The effect of the random instantaneous error at each sampling point is to add a noise component to the signal - referred to as Quantising Noise. There is a simple rule of thumb that can be applied here: The best Signal-to-Noise ratio (S/N) that can be achieved is given simply by S/N (dB) = 6n - 1.75, where 'n' is the number of bits of quantisation and '1.75' is a fiddle factor (and sometimes takes slightly different values in various text books), but S/N is approximately 6n. If a figure of 40dB is taken as 'good communications quality' then 8 bit quantisation, allowing a bit less than 48dB S/N, would be adequate. This is the system adopted on the public telephone network, although in slightly modified form.



The AMBE module.

We can see that for 8000 samples per second, sampling at 8 bits per sample, a total of 8 x 8000 = 64000 bits per second (b/s) is generated [1]. The digital telephone network has a wide enough bandwidth with optical fibres and microwave links to be able to pass 64kb/s directly, but the radio communications link does not have this luxury! At HF we want to be able to pass digital voice over a bandwidth comparable with SSB, ie 3kHz; at VHF, if NBFM is taken for the standard channel width, we can increase this figure to 12kHz, but in order to preserve the enhanced voice quality that good S/N FM can give, more quantising levels should really be used.

Although it is theoretically possible to transmit 64kb/s in a 3000Hz bandwidth, the S/N ratio required to do so with a sufficiently low error rate is very high - around 64dB according to Shannon's Information Theorem - so other techniques have to be adopted to transmit digitised voice signals. Ideally a data rate comparable with the RF bandwidth is wanted for optimum transmission in signal-to-noise ratios that would be just acceptable for poor speech quality, ie around 3000 bits per second for 10-15dB S/N in 3kHz.

VOICE ENCODING

THE TECHNIQUE adopted is to encode the voice to reduce the number of bits per second needed for transmission. There has been a considerable amount of research into various techniques for doing this over the last 10 or so years and some very effective compression schemes are now available. The techniques are too complex to cover in any detail here, but usually involve modelling the human voice tract and coding the various elements, such as voiced and unvoiced sounds. Various schemes such as Code Excited Linear Prediction (CELP) and various other 'ELPs' have come to the fore for voice. As an example, GSM mobile phones use a technique that allows transmission at 13000 bits per second. Whatever technique is used for voice encoding (vocoding), there is usually a trade off to be had between data rate generated and the quality of the resulting speech. Some of the early systems had a very synthetic sounding Dalek-like result, but modern variants provide very much better toll-quality [2] speech. GSM at 13000b/s is an example that has been around for several years now.

Previously, G4GUO had written some voice encoding software for a DSP system based on published algorithms using these techniques, but it was not too successful and

* 15 Noble Road, Hedge End, Sothmpton SO30 0PH. E-mail: actalbot@dera.gov.uk

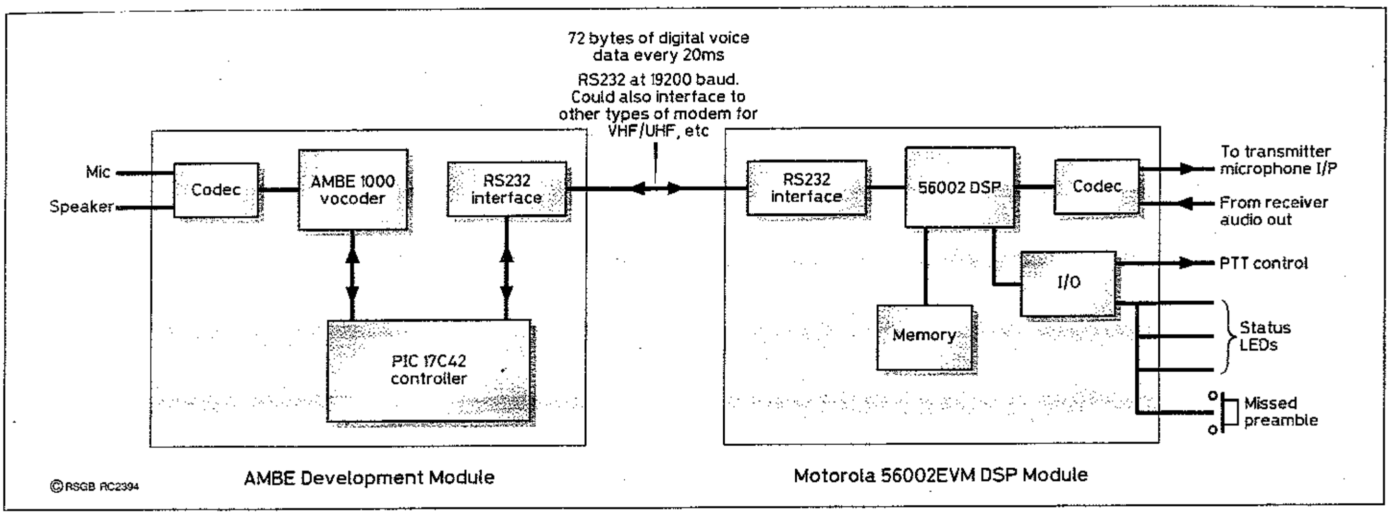


Fig 1: Block diagram of the Digital Voice hardware.

as many of these algorithms are patented using them could run into all sorts of problems or prove expensive in licensing fees! Furthermore, CELP uses a large codebook, too big to fit on any of the low cost DSP development cards.

A comprehensive search of the literature on vocoding techniques threw up yet another new system called Advanced MultiBand Excitation (AMBE) that appeared to offer major improvements over earlier systems. This moves away from the concept of modelling the voice tract, and instead models the spectrum of the signal every 20ms. Not many technical details appear to be available to date as it is still a commercial system, but the results of test programmes show the technique to be better than any of the ELPs, and has thus been adopted for at least one of the new satellite-based mobile phone systems. Much more importantly for us, a single chip solution is available for converting from microphone input to encoded digits, so rather than try to write vocoding DSP software based on published algorithms, it was decided to just buy a chip to do the job.

The AMBE1000™ chip by Digital Voice Systems Inc, (available from Lucent Technologies) implements the whole process and provides the user with an extensive trade off of output data for different levels of link quality, as well as Forward Error Correction (FEC). A range of output data rates from 2400 bits per second all the way up to 9600b/s in 50b/s increments is available, along with a range of user selectable FEC options. Forward Error Correction is essential at HF to overcome burst and CW interference and the effects of multipath - the manufacturers have implemented FEC within the chip that is optimised to the needs of the AMBE technique and in fact claim that "Any other separately implemented FEC technique cannot be as good as our one". Eventually the data rate adopted was 2400b/s of voice data plus 1200b/s of FEC, giving a total of 3600b/s to be transmitted over the RF link. The IC outputs samples every 20ms and can be regarded as a real time system in this sense. Any 20ms samples that get lost just create glitches in the speech that cause minimal disturbance and can often go unnoticed. In fact, during early phases of the experiment, the FEC got left out and tests carried on for some time over our 40m test link before we noticed no error correction was being applied.

VOCODER MODULE

IN USE, the AMBE chip has to be programmed at turn-on to set the operating conditions. The easiest way to do this was to include an on-board PIC microprocessor. G4GUO produced a PCB which includes the AMBE chip, a Coder-Decoder (CODEC) and PIC17C42 controller with other peripheral components. A photograph of this is shown left.

The digitised output samples, at a net overall rate of 3600 bits per second, are sent via an RS232 interface to the modem in packets of nine bytes at a time for each 20ms frame. The data rate used for this part of the link is 19200 baud. If you do the sums on this, it is possible to see there is a lot of spare capacity here for users who want to use the development board for their own purposes. An example of this would be for the inclusion of data and control signals. The PCB was produced as a development tool and thus the controller chip chosen is capable of considerably more processing capability than that needed to just control the AMBE and convert to RS232 format. In fact there is enough processing capacity in the PIC to include encryption / decryption software, but it is left to other users to make use of this capacity! All input/output ports for the PIC are available on connectors.

The AMBE chip is a 100 pin device using the surface mount TQFP package with a lead out spacing of 0.65 mm. Mounting this device caused a few headaches, but use of a good quality magnifier and careful soldering followed by careful inspection resulted in success first time. All the other devices on this board are either conventional surface mount or through hole mounted. **Fig 1** shows a block diagram of the system

RF MODEM CONSIDERATIONS

SO FAR WE HAVE digitised the voice signal and added redundant information to correct potential transmission errors resulting in an overall, but packetised, rate of 3600 bits per second. We now need to turn this data stream into a continuous modulated audio waveform, suitable for feeding to the input of an SSB transmitter for upconversion to the final RF signal for transmission.

There are a number of different ways of modulating an RF waveform for data transmission. Probably the most popular system amongst amateurs is Frequency Shift Keying (FSK), where one audio tone is used for a logic '1' and another tone for logic '0'. In terms of bandwidth utilisation FSK is not very efficient - either a frequency shift significantly wider than the data rate has to be used for reliable copy in noisy conditions, or as the shift is narrowed to preserve bandwidth so errors occur more frequently. The two systems where this trade off is evident is in amateur and commercial RTTY, where 170Hz frequency shift or wider is used for 50-75 b/s data. The converse is HF packet radio, where 200Hz shift is used for 300 b/s with its very poor reliability in noisy conditions. One advantage of FSK is its simplicity of implementation based on filters and phase locked loops, hence its popularity in the past, but with the increased use of Digital Signal Processing (DSP) everywhere, this simplicity is no longer a real issue.

A much better solution is Phase Shift Keying (PSK), where instead of changing the transmission frequency for binary 1s or 0s, the phase is reversed, or effectively the signal is inverted, between 0 and 1 states. It is possible to show theoretically that there is at least a 3dB improvement in S/N versus error rate performance given an 'ideal' demodulator for each modulation, and with simple FSK modems very much better than this is possible in practice. PSK to replace keyboard-to-keyboard RTTY has appeared on the amateur bands recently in the form of PSK31 (see the article in December 1998 and January 1999 *RadCom* by Peter Martinez, G3PLX). For very nearly the same data rate as RTTY, the bandwidth needed has shrunk from around 200Hz to 20Hz, with a corresponding increase in reliability and error rate. By using four phase states 90° apart instead of two, it is even possible to encode two bits at a time without increasing the bandwidth, but this does incur a 3dB penalty as the transmission power is now shared between twice as many bits in a given time. This technique, Quaternary Phase Shift Keying or QPSK, is available in PSK31, where it is included as an option for adding the extra data needed for Forward Error Correction in noisy environments. A properly filtered PSK signal has a bandwidth that is can approach half the baud rate - in fact PSK31 is optimised to do just this - but if not implemented correctly with waveform control and filtering, the bandwidth of the transmitted signal can easily spread alarmingly in a manner analogous to CW key clicks.

Fig 2 shows how raw QPSK modulation by a rectangular data signal would look, showing how the data bits are coded two at a time. This signal as shown in the diagram is unfiltered and would have a very wide frequency bandwidth due to the sharp phase transitions.

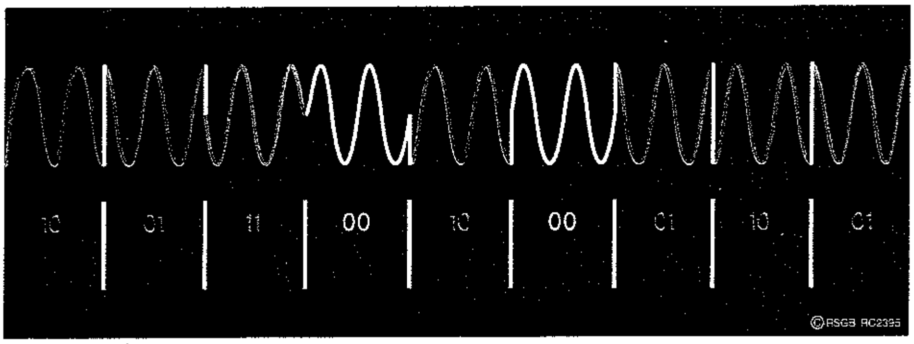


Fig 2: Illustration of an unfiltered QPSK signal, showing how pairs of bits can be coded onto a carrier.

Here we also need to introduce the concept of the Baud rate as opposed to bits per second, which is the rate of information change on the final transmission, or the symbol rate. For BPSK and FSK systems the Baud rate is the same as the bit rate, ie one bit per baud, and the terms are often used interchangeably. Where two bits are transmitted per symbol, as for QPSK, the Baud rate is half that for the given bit rate. It is possible to extend the transmitted bit rate higher by simply increasing the Baud rate to suit the data to be transmitted. For the 3600 b/s needed for the digitised voice experiments, either binary PSK at 3600 baud or QPSK at 1800 baud would be adequate - the QPSK signal at potentially 1800Hz bandwidth could even be transmitted unmodified over SSB radios. However, whilst this technique is ideal for UHF or 'clean' VHF links, there are particular characteristics on a typical HF transmission path that make simple high Baud rate signals very prone to errors and frequently unusable.

REFERENCES

[1] As an aside, consider CD music recording. A sampling rate of 44100Hz is chosen to allow a 20kHz maximum audio frequency; 16 bit quantisation is used to give a dynamic range in greater than 90dB. With two independent channels for stereo, this results in a data rate in excess of 1.4Mb/s.

[2] The term 'toll-quality' comes from America, and means a quality of telephone transmission that people are prepared to pay for. ♦

To be continued...

RadCom Helplines FOR MEMBERS

- AJ Spayne, RS176038, is looking for a wave change switch for an **AR88D**. Also, a service manual and circuit diagram for the **Sony CRF-1** communications receiver. All costs covered. AJ Spayne. Apartado 167, Carcavelos 2776-902, Portugal.
- Dana Jensen collects KW equipment ("virtually unsold and unheard of in North America"), and is looking for a **KW2000E** and **KW600**. E-mail: danaj@home.com
- Jerry, GM4CAZ, would like to know if the **Amateur Radio Mobile Society** still exists. Tel: 0973 937258. E-mail: jclefever @aol.com
- Charles, G4RTV, requires a copy of the operating manual for the **Yaesu FC-902** ATU. All costs covered. G4RTV, QTHR. Tel: 020 8402 2108.
- Richard, G4PRI, requires the loan of the instruction book for the **Katsumi EK-150** electronic squeeze key. All costs covered. G4PRI, QTHR. Tel: 01895 270772.
- Harry, G3MFW, is looking for a manual or the circuit of the **Sony-Tektronix 326** portable oscilloscope, and the 18V power unit for his **Toshiba T2200SX** notebook computer. Costs gladly covered. G3MFW, QTHR. Tel: 01726 73608.
- SP Shackleford, G2HAX, is looking for information on and the circuit diagram of the **Heathkit HM-102** SWR meter. All costs covered. G2HAX, QTHR. Tel: 0118 941 0235.
- Don, G3WDY, would like to borrow a manual for the **Cossor 343** Ganging Oscillator. All expenses paid, and originals speedily returned! G3WDY, QTHR. Tel: 020 8653 4738..
- Sam, G3HVI, is trying to locate a valve type **815** or **QV04-20**, for use in a restored RCA AM transmitter. G3HVI, QTHR. Tel: 01782 393349.
- Ron, G0GHX, is looking for the source of an RF hybrid module **OM631**, as originally fitted in the **Advanced Communications** 6-way TV/FM distribution amplifier. G0GHX, QTHR. Tel: 01202 880194.
- Don, G0ACK, is looking for instructions, information and the circuit diagram of the **Tradiper TE-15** Grid Dip Oscillator. G0ACK, QTHR. Tel: 020 8845 9575.
- Roy, G3JNM, requires a copy of the instructions for assembly and adjustment of the **Hy-Gain 12AVQ** vertical antenna. G3JNM, QTHR. Tel: 01204 843999.
- Alan, GM4IIR, is looking for advice on how to feed a **132ft long wire** antenna. GM4IIR, QTHR.
- Arnold, G8AHE, would like technical help with an **IBM 8513** video monitor. G8AHE, QTHR. Tel: 0121 458 2406.
- Bill, GW3DGT, is looking for a copy of the instruction manual for the **Sharp EL-1615** electronic printing calculator. Also, technical data on the circuitry. W Barrett, 'Stevina', Ludchurch, Narberth, Dyfed, SA67 8JF. Tel: 01834 831369.
- Anthony, M1AVE, is looking for the circuit diagram and alignment details for the **Revco RS2000** AM/FM scanner. Also, any information on the **Anglian 1000** transceiver, especially on the interconnecting cable to the power supply. Tel: 01908 373114.
- VC Whitchurch, G4HSA, would like a copy of the service manual and circuit diagram of the **Philips PM3240** oscilloscope. Also, operating instructions, as in the lid, for the **Wayne Kerr CT492** (mil type) LCR bridge. All costs covered. G4HSA, QTHR. Tel: 01761 414169.
- Ron, G4MNB, is looking for a copy of the manual for the **Yaesu FRG-7** receiver, plus any modifications. All costs reimbursed. G4MNB, QTHR.