DIGITAL SIGNAL PROCESSING ENHANCES A TRACKING SYSTEM

David Spreadbury, Plextek Ltd, UK

Ralph Kanter, TRACKER™ Network (UK) plc, UK

AUTHOR BIOGRAPHICAL NOTES

David Spreadbury graduated from The City University in 1971 with a first class honours degree in Electrical and Electronic Engineering. An experienced hardware designer, he has spent twenty years "doing DSP", in both military and commercial environments using a range of technologies. For the last four years he has been a Senior Technology Consultant at Plextek Ltd., specialising in DSP for digital radio communication. David is married, with two secondary-school aged children. He is actively involved in his local church and enjoys fixing things, reading, origami, computing, and a strictly limited amount of sport.

Ralph Kanter has been Chairman and Chief Executive of TRACKER Network (UK) plc since 1991, and spent three years and £500,000 bringing the product to the UK market, where it was launched in 1993. He was one of the three founders of Britannia Security Group plc in 1983, and before that spent twenty-five years in general management in large corporations including Carreras Rothmans and Thomas Cook, where he was Group Managing Director. Ralph Kanter is also a non-executive director of a number of other public and private companies.

ABSTRACT

The original TRACKER car recovery system is both simple and effective, using a vehicular unit which transmits a location signal in response to a command given over a proprietary VHF radio network. The benefits are a covert, low-cost system, modest radio bandwidth requirements, high location precision, and zero false alarm rate. Drawbacks are that the owner must be aware of the theft, police vehicles cannot be guided into the detection range of the alerting vehicle, and users who require other facilities cannot benefit from this low-cost solution.

More costly systems provide bi-directional communication via networks like the Global System of Mobile Communications (GSM) or INMARSAT. Through the application of Digital Signal Processing (DSP), TRACKER's simple network has now been given bi-directional capability, even though the vehicular Effective Radiated Power (ERP) is only about 30mW, due to the covert antenna.

The immediate uses of this new capability are system monitoring and theft alert. In addition, the integration of a low-cost OEM Global Positioning System (GPS) receiver into the vehicular installation will allow position reporting, which will enhance existing operation and provide many new facilities.

How this is done without compromising TRACKER's simplicity or cost-effectiveness will be described.

1 IMPLEMENTING AN UPLINK

1.1 The Problem

The existing downlink modulation scheme is frequency modulation (FM) by a minimum-shift keying (MSK) 1200 bits per second (bps) data waveform. This waveform uses a single cycle of a pure 1200Hz tone to represent a “0” bit and one and a half cycles of 1800Hz to represent a “1” bit. The peak deviation is ±2kHz, and the resulting spectrum is like that shown in Fig. 1.

The provision of an uplink is complicated by the following considerations:

- the vehicular ERP is low (about 30mW), due to a covert and therefore inefficient (approximately -15dBi) antenna, and
- the network base stations antennae have high visibility to interfering signals, and are often co-sited with antennas from other services.

The net effect of this is to severely reduce the signal to noise ratio of any vehicular signal received by a base station, making a conventional single channel (12.5kHz bandwidth) receiver useless. Receiver noise and probability of interference can be reduced by significantly limiting the bandwidth, but this is easier said than done because, in the interests of economy, the vehicular transmitter frequency is only held to an accuracy of ±1.5kHz.

1.2 The Solution

The solution is to simultaneously modify the modulation format and use a special DSP-based narrow-band receiver, as follows:

1.2.1 Narrow-band Modulation

The modulation is changed by reducing the bit rate significantly, to about 56 bps, and using a fixed deviation for each data bit (i.e. +2kHz for the duration of a “0” bit, -2kHz for the duration of a “1” bit). The digital circuitry of the vehicular unit made this modification very straightforward. Transitions between +2kHz and -2kHz are controlled, using half a cycle of the original 1800Hz waveform, rather than being abrupt. Although this still occupies an overall bandwidth of around 4 kHz, spectral utilisation is limited to two narrow peaks at the two deviation limits, as shown in Fig. 2 (the width of the peaks in this figure is more to do with the spectrum analyser filter settings than the modulation). A little thought will show that since the carrier is being switched between two frequencies, one representing the “0” bits and the other the “1” bits, the result is in fact very similar to two amplitude modulated (AM) signals, 4kHz apart, the lower one being 100% modulated by the data, and the upper one being 100% modulated by the inverse of the data. The spectral width of each of these signals is very low (a few tens of Hz), and thus is a candidate for a low bandwidth receiver.

1.2.2 Narrow-band Receiver

The receiver has to accommodate spreads and variations in the vehicular transmitter’s carrier frequency over a few kHz, so a single narrow-band receiver would be inappropriate. What is required is a large number of narrow-band receivers, evenly spaced (and ideally overlapping) over the whole useable band. With this approach carrier frequency variation becomes a useful feature, because it allows more than one signal to be received at once. For this reason, the vehicular frequency is intentionally dithered by an additional 500Hz, giving a total variation of some ±1.75kHz, so that the overall radio channel utilisation is two 3.5kHz bands, whose centres are 4kHz apart.

Multi-channel reception would be very difficult using analogue circuitry, but is readily achieved with DSP. A conventional single-channel front end produces an output around an audio frequency IF, which is sampled at about 56kHz by an analogue to digital converter (ADC). These samples are passed to a DSP processor for detection and demodulation.

1.2.3 DSP Functions

The samples from the ADC are mixed with digital local oscillators to give two complex baseband signals, 4 kHz apart, corresponding to two components of the transmissions. These in turn are processed with Fast Fourier Transforms (FFTs), which are an efficient way of computing a spectral estimate. Each point of these computed spectra will reflect the energy at a single carrier frequency, and so each AM signal will be visible on one of these spectral points.

The overall implementation results in 256, 33Hz bandwidth receive channels, spaced at 14Hz intervals over each 3.5kHz band, as illustrated in Fig. 3.

Although the upper and lower bands convey the same information (the modulation on the one being the inverse of the other), their separation gives a very useful diversity, increasing the probability of successful reception against narrow band interference.

Each FFT output point is low-pass filtered to 33Hz (to pass the 56 bps modulation but little else) and checked for the presence of the synchronisation pattern which precedes every transmission. Because of the overlapping bands, such detections normally occur in multiple adjacent frequency bands and sampling instants, and taking account of all of these improves both the reliability of the detection and the accuracy of the carrier frequency and bit timing measurements made for the signal.

Demodulation is performed by a number of identical but separate processes. In each, the raw received signal from the ADC is mixed with a local oscillator at the measured signal frequency and low-pass filtered. Detection of the AM proceeds using the measured bit timing.

All this processing is performed in software in an Analog Devices’ SHARC DSP Processor, running at 33MHz, on a very simple short PC card. This has enough capacity to run seven demodulation processes in parallel with the two, 256-channel detection processes. Thus up to 7 signals may be received simultaneously.

2 USES OF AN UPLINK

2.1 Enhancing Existing Facilities

The bi-directional communication of an uplink brings the following improvements to the original system:

2.1.1 Theft Alert

It is now possible for a vehicle to alert the system to its potential theft, by the simple inclusion of a motion detector. Activation is tested against the state of an immobiliser circuit or the presence of the owner’s radio key fob. Unauthorised motion causes an alert uplink message to be transmitted, which allows an owner to be contacted so that a theft can be quickly verified. Note that this procedure, which keeps the owner in the loop, maintains the zero false alarm rate.

2.1.2 System Monitoring

The uplink is also useful for monitoring the status of the vehicular units. This allows not only installation but complete system operation to be checked, which will provide data for both maintaining and improving the system.

2.2 New Facilities

The capacity for the transmission of data from the vehicle opens up the possibility for new applications. as follows:

2.2.1 Driver Assistance

A future application is the transmission of a driver-initiated uplink message, by which a driver may summon roadside assistance. With the current limitation regarding police tracking vehicles having to pass within range this is of limited (but some) use. The inclusion of a vehicular GPS receiver, however (see next section), will make this very useful indeed.

2.2.2 Vehicle Monitoring

There are a number of situations where a vehicle’s owner or the emergency services may benefit from real-time information about the vehicle. Examples are the temperature of a refrigeration unit, the inflation of an airbag, fire detection, etc.

2.2.3 Position Reporting

A commercially available credit-card-sized OEM GPS receiver has been integrated into the vehicular system making convenient use of the previously-developed robust enclosure and supply protection, regulation, and backup circuitry. The Motorola GT Oncore unit was selected from a number on the market because of its low cost and simplicity (which is desirable here because the central controller of the vehicular unit is a microcontroller, not a large microprocessor, and code space and processing power are limited). The position data are output in binary integer form, ideally suited to direct transmission or processing, with the lsb equal to 1 milli-arc-second (mas) (equivalent to a distance of about 3.1 cm at the equator).

To maximise vehicle and backup battery life the GPS receiver, which consumes about 1W, is programmed to take position fixes continuously (every second) when the vehicle's ignition is on, otherwise on demand from the network (such a demand being able to initiate a single or a regular sequence of fixes). An integral rechargeable lithium battery ensures that the GPS receiver remembers the latest fix and satellite data, so as to minimise the time taken for subsequent fixes.

Selective Availability (SA) - an intentional random 29m rms error built into each satellite - results in a positional error of around 50m rms (or <100m for 95% of fixes) for the Standard Positioning Service (civil users)[1]. This error is magnified if the visible satellites have poor geometry, indicated by a large Dilution of Precision (DOP) figure for the fix. These errors are acceptable for localising a vehicle to a motorway or town centre for subsequent tracking, but not for other possible applications. It was therefore seen as desirable to design in the possibility of differential GPS[2] (DGPS), which can reduce the errors by a factor of 10 or so. DGPS works by having a second static GPS unit nearby at an accurately known position. The SA and atmospheric errors will be for the most part common to both receivers, so the differences between their fixes should accurately reflect the real difference between their locations. All that is required, then, is for the surveyed receiver to broadcast the error between its surveyed and observed positions to other nearby mobile receivers.

The catch is that the error in any fix will depend on which satellites were used for that fix, and this will vary from receiver to receiver and fix to fix. Thus the stationary receiver has in reality to compute the range error for each individual satellite, and broadcast this information. Most GPS receivers do incorporate a port specifically designed for the inputting of such correction data, sent in a standard format.

In our application this is a problem because the bandwidth requirements of such a continuous transmission are too high for the network as it currently operates. A more appropriate solution is inverse DGPS (IDGPS), where a positional fix is taken in the usual way and correction is applied subsequently. Here, the network will keep a log (typically half an hour's worth) of the individual satellite range errors and positions, taken every few seconds. These data are conveniently available at the lowest cost from the Motorola VP Oncore 8-channel GPS receiver, although it must be noted that a synchronised pair of such receivers is required for hemispherical (12-satellite) coverage.

The positional message from a vehicle includes the longitude and latitude, time of fix (to the nearest second), type of fix (two- or three-dimensional), and identity of the particular satellites used for the fix. These last three parameters allow the network to compute an individual correction for each message, using either a PC at each base station, or the network's central processing resources. Such an approach is attractive because it has allowed the correction mechanisms to be designed in with minimal impact on the vehicular units, is cheap to implement, and requires only a minor increase in positional message size.

The correction algorithm looks up the range correction necessary for each of the N satellites used at the time of the fix:

$$C = \begin{bmatrix} c_1 \\ c_2 \\ . \\ c_N \end{bmatrix}$$

and the position (relative to the vehicle) for each of the same satellites, in terms of north, east and vertical components of the unit vector:

$$H = \begin{bmatrix} n_1 & e_1 & 1 \\ n_2 & e_2 & 1 \\ . & . & . \\ n_N & e_N & 1 \end{bmatrix} \quad \text{or} \quad H = \begin{bmatrix} n_1 & e_1 & v_1 & 1 \\ n_2 & e_2 & v_2 & 1 \\ n_3 & e_3 & v_3 & 1 \\ . & . & . & . \\ n_N & e_N & v_N & 1 \end{bmatrix}$$

(for two dimensional fixes) (for three dimensional fixes)

From these it computes the least mean squares (LMS) positional correction[3]

$$\text{correction} = \left[\left(H^T \cdot H \right)^{-1} \cdot H^T \right] \cdot C$$

Fully implemented, this correction includes time and, for three-dimensional fixes, height. In this application, only the north and east components of the correction are computed, which reduces the computational load.

To maximise the usefulness of this GPS facility and simultaneously minimise additional radio bandwidth requirements, simple routines which continually process the fix data against geofences (geographic perimeter shapes) have been added to the microcontroller code. A large number of these geofences are predefined in Read-Only Memory (ROM) (circles around seaports, for example, or polygons around national boundaries), and a further limited number can be downloaded over the network into Random Access Memory (RAM), and hence be defined and modified by the user (a polygon around a construction site, for example).

2.2.4 Tracking Vehicle Guidance

The positional information from the GPS receiver can be used to guide a police tracking vehicle directly to within tracking range of a stolen vehicle, thus making more efficient use of limited resources and further reducing the location time.

2.2.5 Progress Monitoring

Positional information will enable the progress of industrial vehicles to be monitored. The improved accuracy of IDGPS will allow the monitoring of the routes of hazardous loads.

2.2.6 Additional Theft Alert

The geofences provide an additional mechanism to generate an early alert to possible theft, even with an authorised driver, by initiating the transmission of a warning message whenever the vehicle moves into a forbidden or suspect area. Note that geofence calculations performed inside the vehicle are subject to the basic non-differential GPS errors; this is quite adequate for an alert mechanism.

3 IMPLEMENTATION STATUS

Following a successful trial phase (with ranges of up to 70 miles being achieved), the uplink described in this paper is now in every day use.

The GPS receiver integration and vehicular software are complete, and trials are in progress. The other features mentioned are in various stages of development.

4 CONCLUSIONS

State-of-the-art DSP techniques have been described which have enabled a simple unidirectional VHF network to be given a bidirectional capability, in difficult circumstances. This has been achieved at low cost to both the vehicular installation and the network. The approach, which has been proven, will pave the way for additional features which hitherto have only been available from significantly more expensive networks.

5 REFERENCES

1. “GPS Navstar User’s Overview”, Fifth Edition, ARINC Research Corporation, Los Angeles, March 1991, pp 42, 56-59.
2. “Navstar GPS User Equipment Introduction”, Public Release Version, NAVSTAR-GPS Joint Program Office, Los Angeles Air Force Base, February 1991, chapter 11.
3. Kaplan, Elliott D: “Understanding GPS: principles and applications”, Artech House, 1996, ISBN 0-89006-793-7, pp 519-522.

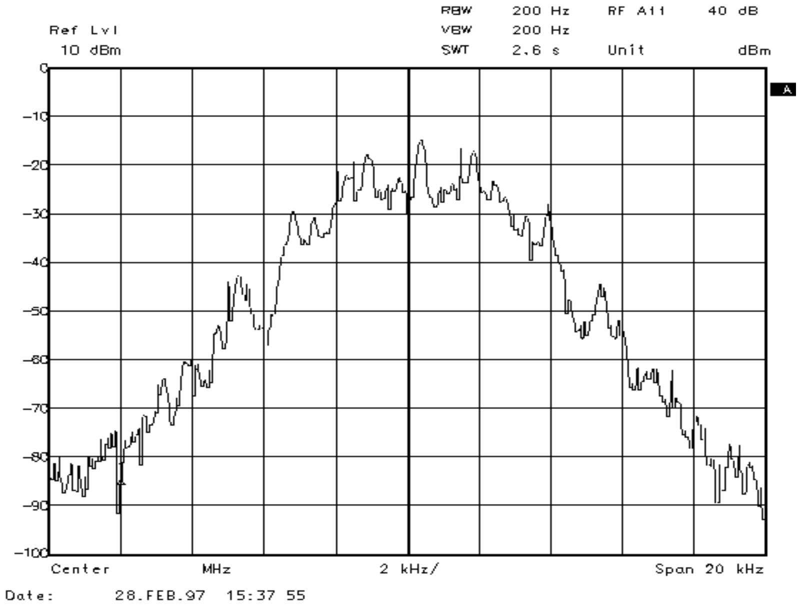

FIGURE 1: Spectrum of 1200 bps Modulation

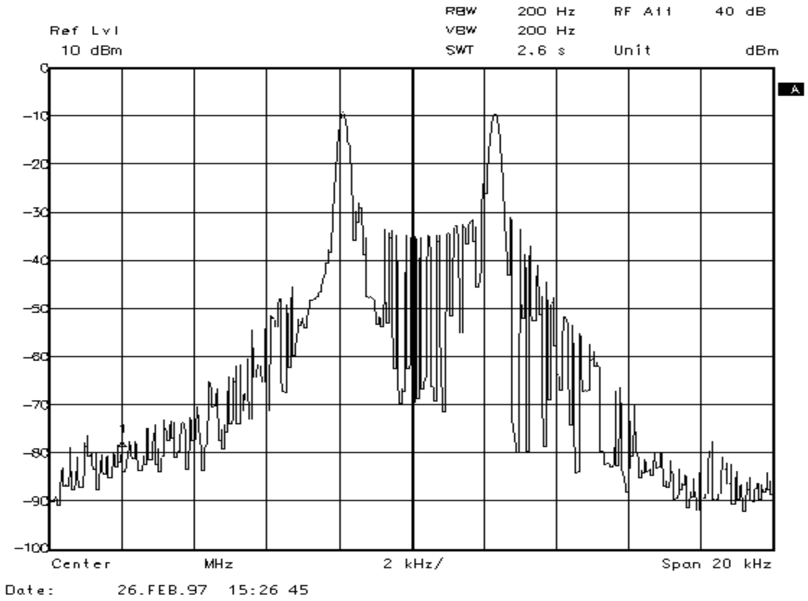

FIGURE 2: Spectrum of 56 bps Modulation

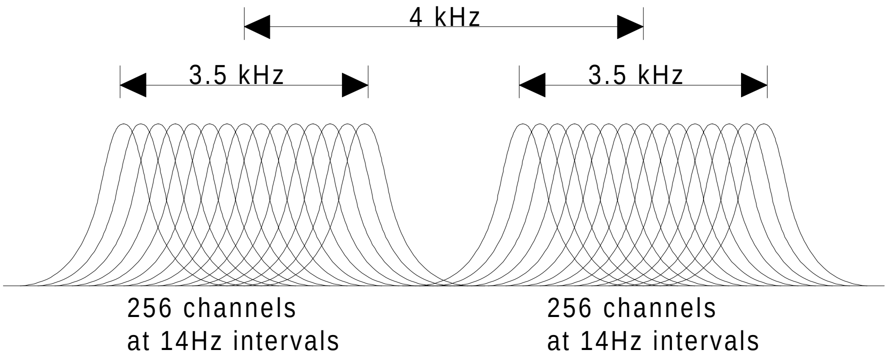

FIGURE 3: DSP Narrow-Band Receiver Response