Assisted GPS

Using Cellular Telephone Networks for GPS Anywhere

Rod Bryant, SigNav Pty. Ltd.

THE FIRST TWO ARTICLES TO APPEAR IN THE INNOVATION COLUMN IN *GPS WORLD* more than 15 years ago were entitled "GPS: A Multipurpose System" and "The Limitations of GPS." In the first article, GPS was trumpeted as a revolutionary positioning and navigation technique that could be used in all sorts of unexpected ways. This premise has stood the test of time with new uses for GPS still being discovered. The second article reminded readers that GPS may not be a panacea for all of our positioning needs — that there are some situations in which GPS fails us. In particular, it was noted that GPS signals are "blocked" by buildings, making indoor use of GPS impossible. But what was impossible 15 years ago is possible today. New designs have greatly improved the sensitivity of GPS receivers so they can make code-phase measurements even on the severely attenuated signals inside buildings. And if the signal is too weak for the receiver to extract the satellite navigation message itself, the necessary data can be sent to the receiver using a cellular telephone network, which also can supply timing information to help the GPS receiver acquire signals more quickly. In this month's column, we will investigate how this so-called "assisted GPS" works and why we can now say that GPS works (virtually) anywhere. — R.B.L.

The integration of GPS into cellular telephones enables a potentially vast array of new applications ranging from consumer gimmicks through efficiency multipliers for enterprises to lifesaving safety and security applications. In the United States, the Enhanced-911 regulations have been and remain a major catalyst for this deployment. In Europe, the commercial potential of location-based services (LBS) is driving it. Regardless of the drivers, the technology convergence is happening with an ever-increasing momentum.

These new applications and the cell-phone environment itself, however, pose significant challenges for the GPS community. They demand GPS solutions that can be implemented in tiny spaces at extremely low cost in extremely high volumes, operate reliably in a much broader range of environments than was hitherto considered possible, acquire signals in seconds under extreme conditions and, for some applications, do so without the aid of stored orbital data.

These problems and the availability of cellular communications itself spawned the concept of assisted GPS (AGPS) in which the network assists the GPS receiver to perform its various functions. This article reports on AGPS developments leveraging cellular telephone networks to help acquire and deliver accurate GPS fixes from anywhere, anytime.

System Considerations

In developing a user system for AGPS, several factors must be considered, including the type of assistance to be provided by the network, the type of cell-phone network and the corresponding AGPS standards, and environmental factors such as radio frequency (RF) compatibility between the GPS module and the host platform.

Benefits of Assistance. There are many types of assistance that can be provided by the network to the GPS receiver. The receiver could be a fully functional receiver capable of selecting satellites, acquiring signals, achieving time synchronization, extracting data, performing measurements, and computing its own navigation solution. Nevertheless, its acquisition speed can be enhanced through the provision of assistance. Furthermore, it can avoid spending time to extract all of the required data from the satellite signals if most of this is supplied by the network.

Signal Levels. More importantly, the range of signal levels at which the receiver can operate can be greatly increased if the receiver is relieved of the requirement to extract the 50 bits-per-second navigation data stream that is modulated onto the signals. This data cannot be extracted in a timely manner (or at all in many cases) if the received signal power is below about –172 dBW (–142 dBm) but code-phase measurements can still be made for much weaker signals than this. Since the navigation data is not location specific, it can be supplied by a remote receiver that has a clear line of sight to the same satellites.

The data supplied by the network in this way can include ephemeris coefficients, almanac coefficients, satellite health data, satellite clock error coefficients, atmospheric error coefficients, and so on. Additionally, excerpts from the data sequence can be supplied to facilitate coherent integration over periods much longer than a navigation data bit interval.

MS/UE Assisted. In principle, the receiver can be simplified if some of the computation, such as the navigation solution, is performed by the network. This form of operation is referred to as mobile station–assisted (MS-Assisted) or user equipment–assisted (UE-Assisted) in contrast to solutions performed in the user's equipment: MS-Based or UE-Based.

The advantage in terms of simplifying the handset is questionable given the computational capacity of modern cell-phone and GPS chipsets. Furthermore, if the location information is required in the handset, this advantage is offset by the need to communicate back and forth. Nevertheless, such architectures are commonplace.

The network has coarse information about the location of a receiver embedded in a cell phone. If it supplies this to the receiver, along with its uncertainty, the receiver then can use this to determine which satellites to search for, restrict its search ranges, and initialize its navigation solution. The uncertainty typically is supplied in the form of either the size and orientation of the semi-major and semi-minor axes of an ellipsoid of uncertainty, or the horizontal radius and vertical height of a cylinder of uncertainty.

Time Assistance. The network also can supply time and its uncertainty. This may be supplied precisely using hardware in the handset, or much more coarsely over the air from a server. If it is sufficiently precise (for example, to within a few microseconds), it can be used to restrict the search range for the absolute code phases of the satellite signals. If it is less precise (for example, only to within many microseconds to a few seconds), it can be used to restrict the search range for the relative code phases of the satellite signals once an initial signal has been acquired.

If the time assistance is fairly precise (to a few milliseconds or better), the handset position errors resulting from the uncertainty in the satellite positions corresponding to the time uncertainty will be small enough (a few meters) to be tolerated in most cases.

However, if it is coarser than this, the navigation solution also will have to solve for the time error. Note that the receiver may only have code-phase measurements rather than full pseudorange measurements as used in a normal receiver — requiring a procedure to resolve the one-millisecond code-phase ambiguities and a special algorithm to solve for the time error.

Additional Assistance. Alternatively to supplying position, time, and satellite-derived data, the network can derive and provide other assistance including approximate code phases at a certain instant in time, with the corresponding uncertainties, as well as the Doppler offsets with their uncertainties.

Assistance can be supplied in the "user plane" or in the "control plane."

In the latter case, assistance is supplied via communication over the signaling channels from a server integrated into the network infrastructure. The standards discussed in a later section relate to this form of assistance.

In the former case, the assistance is supplied from a user server (typically Web-based), using standard communications channels such as Short Messaging Service (SMS) and General Packet Radio Service (GPRS) over Global System for Mobile Communications (GSM) networks, or Single Carrier Radio Transmission Technology (1xRTT) over code division multiple access (CDMA) networks.

CDMA vs. GSM

There are technical differences between GSM and CDMA cellular technologies that impact AGPS implementation in these networks. Equally importantly, the two communities have evolved different AGPS standards discussed in the next section.

Precise Timing. The first technical difference relates to the fact that precise timing is fundamental to CDMA operation. The handset synchronizes to the communications code very precisely (that is, well below the microsecond level). Using hardware (for example, a pulse and message), precise network time can be transferred to the GPS receiver subsystem. Of course, this time will contain an error equal to the communication latency between the network and the handset, but it is more than adequate as precise time assistance for the AGPS purposes described in the previous section.

GSM does not use spread spectrum codes, and hence this form of precise time assistance is not intrinsically available in a GSM network. To deliver precise time assistance (with uncertainties of 5 or 10 microseconds), GSM networks have to be augmented. Not surprisingly, few network operators are keen to roll out additional infrastructure for this purpose, and GSM deployment of AGPS typically is required to operate with coarse time assistance.

Time Slots. Another technical difference relates to the fact that GSM uses short time slots, so each handset communicates in frequent short bursts. CDMA handsets, on the other hand, communicate using much longer bursts. This impacts on the forms of cell-phone interference mitigation techniques that can be employed by AGPS solutions in the two cases as will be discussed in a later section.

Cell Sizes. A third technical difference relates to the fact that the GSM technology has a limitation on its cell sizes of around 35 kilometers in radius. CDMA cell sizes, on the other hand, are only limited by transmission power and relevant standards (such as CDMA code-phase search ranges). Hence, they can be much larger. If the coarse location assistance is derived from the cell location, its uncertainty can be much larger in a CDMA system than in a GSM system. In practice, however, this is not a significant factor because the most-demanding AGPS environments tend to be in inner-cities where cell sizes can be limited to a few kilometers in radius.

AGPS Standards

Both the CDMA and the GSM communities have developed standards for control plane AGPS messaging (TIA/EIA/IS-801-1, 3GPP2 C.S0022-0-1, 3GPP TS 25.331) and for minimum operational performance of AGPS handsets (TIA 916, 3GPP2 C.P9004-0, 3GPP TS 25.171 V6.0.0). There is considerable similarity between the assistance fields included in the two protocols. The minimum performance standards are measured in both cases using five separate statistical tests.

The five tests are of sensitivity, nominal accuracy, dynamic range, multipath scenario, and moving scenario with periodic update. The nominal accuracy tests are for static accuracy under typical signal strength conditions rather than weak signal conditions and with no multipath present. Performance in the presence of multipath is tested separately, as are the performances under weak signal conditions and under typical land-based dynamic conditions.

One difference between the two performance standards is the handset must respond within 16 seconds in the CDMA case but has 20 seconds to respond in the GSM case. In both cases the required sensitivity is –147 dBm, and the horizontal positioning accuracy, although defined differently, is similarly around 30 meters.

Another difference is the GSM standard allows for either precise or coarse time assistance. When only coarse time assistance is provided, the sensitivity test is conducted with one satellite at –142 dBm. This is a recognition there is a performance penalty for not providing precise time assistance.

The other main difference is that, in the case of MS-Assisted (UE-Assisted) operation, the CDMA standard calls for the absolute code-phase accuracy to be tested, whereas the GSM standard calls for the location to be computed in accordance with a defined algorithm and for the accuracy of the result to be tested. This is more significant than it seems because the code-phase test is of absolute code-phase accuracy rather than relative code-phase accuracy. To pass this test, the time assistance must be used to determine the measurement instant with nanosecond precision.

The reason for this requirement is the location server can combine GPS and CDMA measurements in performing a hybrid fix only if the absolute GPS code-phase measurements are known for a precise time (according to the CDMA handset's local clock).

These standards have emerged from complex techno-political negotiations between network operators, handset manufacturers, technology providers, and semiconductor manufacturers. They represent negotiated compromises between these various groups rather than a true consensus as to real-world requirements. In particular, the author considers the sensitivity requirement to be lacking in stringency. For reliable positioning under most indoor conditions, sensitivity of at least –150 dBm is essential and better than –153 dBm is desirable. Sensitivity of better than –185dBm is ideal.

Cell-Phone Interference

One of the technical problems facing the GPS cell-phone developer is the interference to GPS reception from the very strong cellular transmissions of the handset. This is an even more serious issue given that the GPS front-end and antenna performance typically is compromised as a result of the severe physical constraints on the design. It is further exacerbated by the need for extreme sensitivity.

The ideal solution to this problem is to provide sufficient filtering in the GPS RF path to permit concurrent operation of the receiver and handset transmitter. However, the significant benefits that flow from such an approach come at a cost. In particular, the more complex RF design results in additions to the bill of materials that add cost and space.

Furthermore, it inevitably degrades the front-end noise figure of the GPS receiver, thereby offsetting performance gains relative to alternative approaches.

CDMA Mode Switching. In the case of a CDMA handset, the main alternative is to switch modes between GPS and the handset. The two subsystems cooperate so the receiver "listens" only during timeslots when the handset is not permitted to transmit.

For example, during the 16-second response period, the transmitter may only be able to transmit for a fraction of the time in bursts of a few hundred milliseconds. While this process allows roaming to continue, it may well represent an unacceptable limitation on speech communication during that period. Meanwhile, the imposition of antenna switching constraints on GPS signal processing also is significant as integration periods must be sized to fit within these constraints.

GSM Signal Blanking. For GSM, the problem is less extreme because the technology uses timeslots of only a few milliseconds in width. Signal blanking can be used during these slots, thereby efficiently avoiding the need for cooperation between the software of the two subsystems.

Nevertheless, the blanking itself eats into the GPS integration periods, thereby degrading sensitivity and/or increasing typical acquisition times.

Receiver Architectures

In tailoring a GPS receiver for embedded AGPS applications, several factors need to be considered including correlator design, memory requirements, and microprocessor control issues.

AGPS receivers need to acquire weak signals quickly. To meet the demands of the market and exceed the demands of the AGPS standards, more-sophisticated baseband hardware is needed than was required of conventional GPS receivers of only a few years ago. Essentially, this means many more correlator taps or "fingers." How those fingers are organized and the type of signal processing employed are critical factors.

Limited Coherent Integration. **Figure 1** illustrates a common search engine

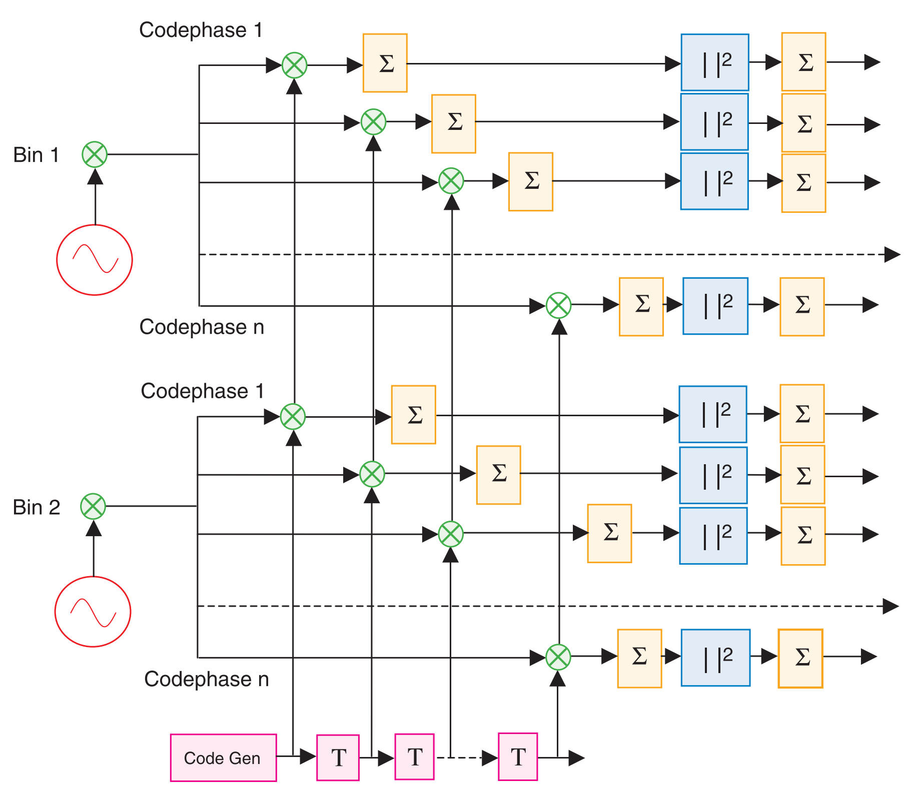


Figure 1 A two-bin hardware search engine. The first set of mixers mix the final local oscillator signals with the incoming signal to downconvert the selected satellite signal to near baseband. The first set of summing blocks perform coherent integration on the downconverted, despread signal. The squaring blocks compute the magnitudes of the complex integrals resulting from the coherent integrations. The second set of summing blocks perform non-coherent integration on the squarer outputs. A shift register produces multiply delayed versions of the locally generated pseudorandom noise code.

architecture. This design performs multiple rounds of coherent integration for each of the *n* fingers per frequency bin, and integrates the results non-coherently. The architecture lends itself to efficient hardware mechanization through reuse of the arithmetic elements. Chips have been produced using this design to incorporate multiple bins and 20,000 fingers or more. Other designs, such as u-Nav Microelectronics' uN8130 baseband chip (with which the author is very familiar), combines a 2,048 finger × four-bin search engine with 12 four-finger correlators.

The first limitation of this hardware search engine is the coherent integration period is limited to much less than a navigation message bit. If not, the probability of bit transitions occurring within integration periods will be high, and excessive random losses will result. The effect of this limitation is to ensure the squaring losses prior to the non-coherent integration are relatively large. The end result is relatively long overall integration periods are required to achieve the desired sensitivity.

The second limitation of this approach is that, when it is possible to constrain the search to a small range of code phases, the rest of the fingers are effectively wasted. When precise time assistance is available this means most of the potential of the search engine hardware is wasted all of the time. When only coarse time assistance is available, it means the full potential of the hardware is being utilized for acquiring the first satellite signal — but, again, most of its capacity is wasted when acquiring subsequent satellite signals.

Flexible Design. Figure 2 illustrates an alternative organization designed to address these issues. In this case, the signal processing of each bin is much more sophisticated, and hence the entire search engine no longer lends itself to hardware mechanization. Instead, it is best implemented as a set of correlator channels each with multiple fingers. By using multiple channels together, one or more larger search engines can be built up as needed to span the required code-phase search range.

This arrangement draws on the patented subATTO signal processing technology (see side bar), which facilitates coherent integration over much longer intervals than a bit period. It results in shorter integration periods being used to achieve the same sensitivity. Since this more-flexible architecture also allows all of the hardware resources to be effective all of the time during acquisition, it results in far more cost-effective use of hardware to achieve the required sensitivity and acquisition time.

This approach has been employed using two very different hardware architectures. In the uN8130 baseband processor, a modestly dimensioned hardware search engine is used in tandem with 12 correlator channels comprising four fingers each. The correlators are used for both acquisition and tracking while the search engine performs the more difficult task in acquiring the first satellite and some of the subsequent satellites. Whenever the search engine acquires a signal, it passes it to one of the correlators. This solution has been adapted successfully for both CDMA and GSM operation, with performance well in excess of the standards.

In another example, flexible correlator hardware resources have been incorporated into a chip designed to operate with a host processor. subATTO processing takes place on the host. A search strategy was devised that keeps the hardware working close to its maximum potential throughout the acquisition phase. The resulting performance is well in excess of the standards yet again with minimal hardware costs. This approach demands much more general purpose processing capacity but much less dedicated hardware support.

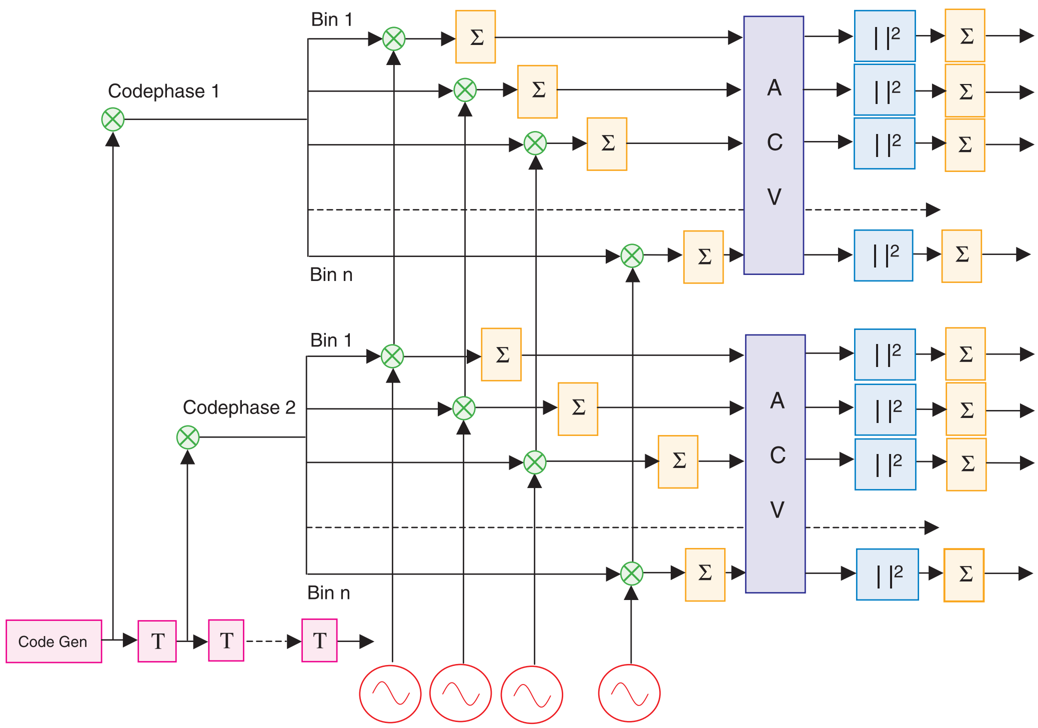
Acquisition Strategies. We have developed a range of strategies to suit GSM and CDMA standards–based assistance schemes and the hardware architectures previously described. These involve the signal processing schemes described above as well as bit-synchronous signal processing schemes (when bit synchronization is feasible) using both correlators and search engines. We are also working with our chipset partners to optimize their hardware architectures to suit advanced acquisition strategies to improve sensitivity and acquisition time. One example is the use of multiple smaller search engines combined with correlators. This is an ideal compromise when general purpose processing capacity is limited.

Memory and Hosting Constraints. In developing the technology, we were required to accommodate a range of memory and hosting constraints. Using the u-Nav chipset, we have implemented an MS-Assisted solution in which the firmware is serially booted into the on-chip static random access memory (SRAM) from the cell baseband chip. This required all of the AGPS firmware to run in 64 Kbytes of program memory and 60 Kbytes of data memory. We also have implemented MS-Based and assisted conventional solutions utilizing external flash memory for the program but internal data memory only.

Utilizing the flexible correlator hardware mentioned earlier, a firmware solution running on a host processor performs all of the subATTO processing to produce codephase and Doppler measurements or location.

The firmware supports MS-Assisted, MS-Based, and multi-mode operation. It runs on an ARM 9 microprocessor but can easily be ported to others.

It has been uniquely structured to facilitate sharing of the processing resources with foreign, high priority applications as might be found on a cell baseband processor. The firmware adapts gracefully to the loss of processing capacity and loss of signal samples. It is independent of the platform's real-time operating system (RTOS) and can be collapsed into a single task.



▲ **Figure 2** A two-code-phase subATTO search engine. In addition to the operational blocks described in Figure 1, the ACV (autoconvolution) block constitutes a proprietary algorithm utilizing multiple fast Fourier transform (FFT) bins. The FFT/ACV/squarer process provides improved sensitivity for the same overall integration period compared to the technique of Figure 1.

Conclusions

GPS operation can be enhanced in its performance through the provision of assistance data wirelessly.

This assistance can be supplied via the user plane or control plane. It can consist of satellite data gathered from remote GPS receivers with direct line of sight to the satellites, coarse receiver position, precise or coarse time, or more specific data derived from these primary elements.

The wireless network may also perform the location solution, although the benefit of this is debatable. The use of assistance can result in much faster acquisition of weaker signals, and can facilitate navigation solutions that would not otherwise be possible.

The benefit of the time assistance is dependent in a complicated way on its uncertainty. Both the search strategy and the navigation solution have to be adapted to the level of this uncertainty.

CDMA inherently facilitates precise time assistance. GSM networks, however, have to be augmented to provide it and generally will not be.

On the other hand, CDMA cell sizes can be much larger than GSM cell sizes, resulting, in principle, in slower signal acquisition and/or a need for more GPS processing capacity. However, this is not a significant factor in practice.

GSM uses short time slots for transmission and this facilitates interference mitigation via signal blanking. A CDMA handset may employ "antenna switching," but with potentially serious impact on both GPS and voice communication. RF design for concurrent operation is far preferable.

There are some differences between the standards developed for CDMA and GSM but, in general, the assistance supplied and the performance requirements are quite similar. However, the sensitivity required by the standards is well under that required to provide reliable performance under a range of indoor and urban canyon environments.

Hardware-mechanized search engines are useful but typically are not flexible enough to permit the potential of the hardware to be fully realized. subATTO signal processing allows similar performance to be achieved with far less dedicated hardware support, but requires much more general purpose processing capacity. A hybrid of multiple smaller search engines and correlators provide an ideal compromise where such processing capacity is not available.

AGPS solutions require a range of different hardware and software solutions. Our team has experience with hardware employing both correlator fingers and search engines, and with flexible correlator hardware designed to work with a host processor. Our firmware solutions have ranged from MS-Assisted firmware running in 64 KBytes of internal SRAM to taskless, RTOS–independent multi-mode firmware running on a host processor that also supports high priority foreign tasks.

Acknowledgments

The entire engineering team at SigNav contributed to this work through their innovation and tireless dedication. In particular, I wish to thank Eamonn Glennon, who has made major algorithmic contributions. Our colleagues at u-Nav and its partners and customers contributed to our growing collection of insights, as have those of our other partner who cannot yet be named. This article is based on the paper "Lessons Learnt in Assisted GPS" presented at GNSS 2004, the 2004 International Symposium on GNSS/GPS, held in Sydney, Australia, December 6-8, 2004.

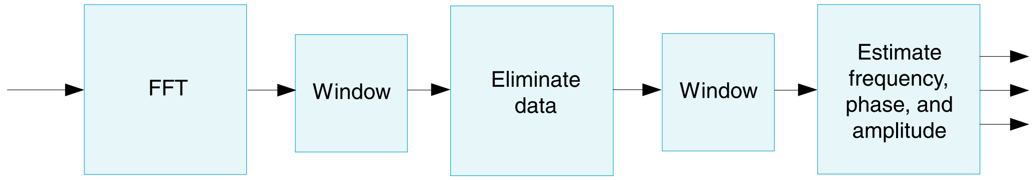


Figure 3 Core signal-processing scheme. The fast Fourier transform (FFT) block is equivalent to the ensemble of local oscillators, second mixers, and first summers of Figure 2, while the windowing and eliminate data blocks correspond to the ACV (autoconvolution) block in Figure 2. Various estimation algorithms are employed depending on the implementation.

Biography

ROD BRYANT is the chief executive officer and chief technology officer of SigNav Pty. Ltd. in Canberra, Australia. Until 1994, Bryant worked at Auspace Ltd. for eight years in a variety of engineering, management, and business development roles associated with satellite engineering. He has a Ph.D. in real-time ultrasonic imaging from the University of Adelaide.

"Innovation" is a regular column featuring discussions about recent advances in GPS technology and its applications as well as the fundamentals of GPS positioning. The column is coordinated by **RICHARD LANGLEY** of the Department of Geodesy and Geomatics Engineering at the University of New Brunswick, who appreciates receiving your comments and topic suggestions. To contact him, see the "Columnists" section on page 6 of this issue.

subATTO Tracking Solution

We have called our solution subATTO because it provides for acquisition and tracking of GPS signals weaker than an attowatt (10^{-18} watts or –180 dBW). This level of sensitivity is essential to provide reliable positioning inside buildings, multi-story car parks, and in urban canyons. The sensitivity achieved is highly competitive with respect to other AGPS schemes reported.

One important aspect of the solution is the subATTO algorithms are integrated with normal receiver firmware. Under typical conditions, the receiver behaves like a normal, fully autonomous receiver that acquires the navigation message data in the standard way. The subATTO solution takes advantage of this.

In the early 1990s, we developed a patented phase coherent open-loop (PCOL) carrier processing scheme providing faster reacquisition and far greater dynamic resistance than conventional tracking loop techniques. Though considerably refined since the awarding of a patent in 1991, this scheme was the basis on which we built the subATTO signal processing solution.

To achieve sensitivity to around –185 dBW (equivalent to a carrier-to-noise density ratio of 19 dB–Hz) or below and be capable of operating in a severe fading environment, it is essential that the signal processing technique require no prior knowledge of the navigation message data. PCOL permits code and carrier tracking completely independently of the data extraction, and thus was an ideal basis for the development of subATTO.

The secret of this approach is to eliminate the data modulation in a way that minimizes the loss associated with the non-linear processing required.

As illustrated in **Figure 3**, this is achieved by performing band-limited non-linear operations in the frequency domain. Note that Figure 3 is indicative only. Although each output bin is derived from a limited portion of the input bandwidth, this is actually done in such a way that the post-correlation bandwidth is preserved. Hence one frequency search step typically covers 1 kHz depending on the post-correlation sampling rate. After eliminating the data modulation, the carrier phase and frequency is estimated and these values are used in obtaining the early and late amplitude estimates required for code tracking.

Refining this scheme to cope with carrier-to-noise density ratios as low as 19 dB–Hz requires the use of long fast Fourier transforms (FFTs) and the addition of a further layer of complexity to the signal-processing algorithms in the form of various ensemble averaging techniques.

Another aspect of the subATTO technology relates to the algorithms we have developed for resolving code phase ambiguity and estimating time without using bit synchronization or the navigation data. The last piece of the jig-saw puzzle is the suite of acquisition strategies we have developed to provide optimal trade-offs of time-to-first-fix versus sensitivity under various assistance scenarios using various hardware facilities.

Further Reading

For further information on subATTO, see

"A GPS Receiver Optimized for Land Vehicles" by R.C. Bryant in *Proceedings of ION GPS-93*, the 6th International Technical Meeting of the Satellite Division of The Institute of Navigation, Salt Lake City, Utah, September 22-24, 1993, Volume II, pp. 1599-1606.

"subATTO Indoor GPS – Pitfalls, Solutions and Performance Using a Conventional Correlator" by R. Bryant in *GPS Solutions*, Vol. 6, No. 3, 2002, pp. 138-148.

For previous *GPS World* articles on assisted GPS, see

"Wireless-Assisted GPS: Keeping Time With Mobiles" by J.T. Syrjärinne in *GPS World*, Vol. 12, No. 1, January 2001, pp. 22-31.

"GPS Reference Networks' New Role: Providing Continuity and Coverage" by P. Enge, R. Fan, and A. Tiwari in *GPS World*, Vol. 12, No. 7, July 2001, pp. 38-45.

"Indoor GPS: The No-chip Challenge" by F. van Diggelen and C. Abraham in *GPS World*, Vol. 12, No. 9, September 2001, pp. 50-58.

"Assisted GPS: A Low-Infrastructure Approach" by J. LaMance, J. DeSalas, and J. Järvinen in *GPS World*, Vol. 13, No. 3, March 2002, pp. 46-51.

INDEX

ADVERTISERS INDEX

COMPANY INDEX