

INTELLIGENT

I

TRANSPORTATION

T

SYSTEMS

S

Vehicle Location and Navigation Systems

Yilin Zhao

VEHICLE LOCATION AND NAVIGATION SYSTEMS

Yilin Zhao

Artech House, Inc.
Boston • London

Library of Congress Cataloging-in-Publication Data

Zhao, Yilin.

Vehicle location and navigation systems /Yilin Zhao.

p. cm.

Includes bibliographical references and index.

ISBN 0-89006-861-5 (alk. paper)

1. Intelligent Vehicle Highway Systems. 2. Motor vehicles—Automatic location systems. 3. Electronics in navigation.

I. Title.

TE228.3Z45 1997

629.2'7—dc21

97-4200

CIP

British Library Cataloguing in Publication Data

Zhao, Yilin

Vehicle location and navigation systems

1. Intelligent Vehicle Highway Systems 2. Motor vehicles—Automatic location systems

I. Title

625.7'94

ISBN 0-89006-861-5

Cover design by Jennifer Makower

© 1997 ARTECH HOUSE, INC.

685 Canton Street

Norwood, MA 02062

All rights reserved. Printed and bound in the United States of America. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Artech House cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

International Standard Book Number: 0-89006-861-5

Library of Congress Catalog Card Number: 97-4200

10 9 8 7 6 5 4 3

To everyone who

and minimum points V_x and V_y to reconstruct the circle center dynamically. Another method is to introduce a correction table. The correction table contains the correction information for any heading. For instance, in Figure 3.13 the compass has an error of -4 degrees at 100 degrees, thus the table will contain a $+4$ -degree correction in the cell for a 100 -degree heading. Adding the correction to the compass measurement will eliminate the error. Independent sources of heading information (i.e., GPS or a digital map) can be used for determining the correction and generating the table. These heading sources are often more accurate than the compass but are not available on a continuous basis.

When the vehicle is traveling over a hilly road, the compass plane will not be parallel to the plane of the Earth's surface. The compass measures only the projection of the vector components. This type of error can be detected and removed via a check on the magnitude of the magnetic field.

3.4.2 Global Positioning System

The GPS is a satellite-based radio navigation system. It provides a practical and affordable means of determining position, velocity, and time around the globe. GPS was designed and paid for by the U.S. Department of Defense (DOD). Civilian access is guaranteed through an agreement between the DOD and the Department of Transportation (DOT).

GPS consists of three main parts: the space segment (satellites), the user segment (receivers), and the control segment (management and control). From the location

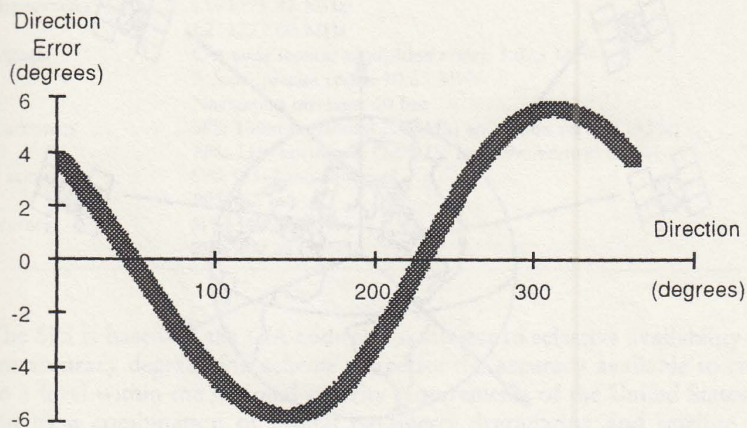


Figure 3.13 Errors due to the shift of the vehicle magnetic field center.

and navigation point of view, only the first two segments are of interest. For more detailed descriptions of all the segments, as well as various theoretical and practical aspects of GPS, readers may consult [17–19].

The observation of at least four satellites simultaneously will permit determination of the 3D coordinates of the receiver (Figure 3.14) and the time offset between the receiver and GPS time. The 3D coordinates can be computed in an ECEF Cartesian coordinate frame or in a geodetic coordinate system such as WGS 84. Using transformation equations, positions can be computed in other geodetic frames or mapping planes (see Section 2.3).

The GPS constellation consists of 24 satellites arranged in six orbital planes with 4 satellites per orbital plane as shown in Figure 3.15. It is designed to provide worldwide coverage 24 hours per day. The characteristics of GPS are summarized in Table 3.3 (some items in Table 3.3 are from [20]). The design life span of each GPS satellite is 7.5 years. The on-board time and frequency standards are the most likely functions to fail, so each satellite carries four atomic clocks on board.

In Table 3.3, SPS stands for Standard Positioning Service, PPS stands for Precise Positioning Service, and 2dRMS stands for two times the root mean square distance error (twice the standard deviation). The term 100m 2dRMS means that 95% of the time the horizontal positions will be within a circle of a 100m radius. Both the C/A code and P code are pseudorandom noise (PRN) ranging codes. PRN is a sequence of digital 1's and 0's that appear to be randomly distributed like noise, but which can be exactly reproduced. The navigation message broadcast includes satellite orbital position data and satellite health status.

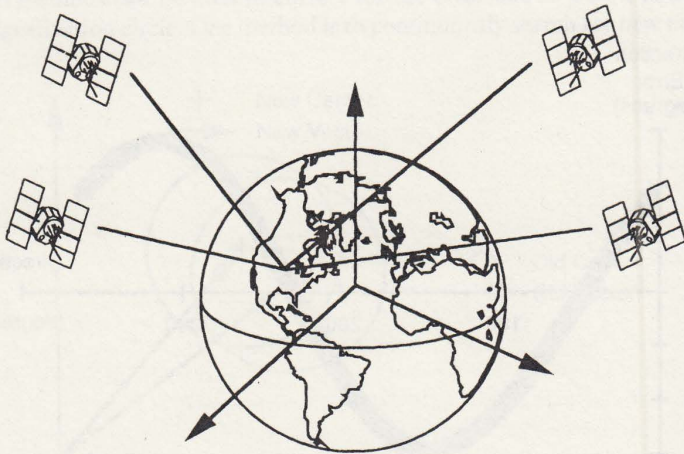


Figure 3.14 ECEF Cartesian coordinate system for GPS.

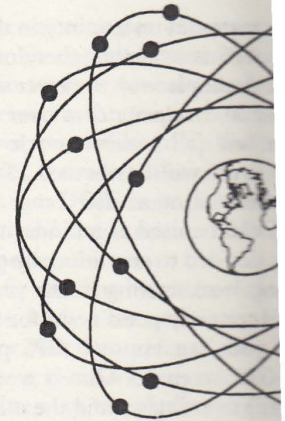


Figure 3.15 Satellite constellation (space segment).

Table 3.3
GPS Characteristics

Item	
Satellites	24 satellites broadcast signals
Orbits	Six planes, at 55-degree inclination, with 4 satellites at 20,231-km altitude
Carrier frequencies	L1: 1575.42 MHz; L2: 1227.60 MHz
Digital signals	C/A code (coarse acquisition) P code (precise code): 10.23 MHz
Position accuracy	Navigation message: 50 bps SPS: 100m horizontal (2dRMS) PPS: 21m horizontal (2dRMS)
Velocity accuracy	SPS: 0.5–2 m/s observed PPS: 0.2 m/s
Time accuracy	SPS: 340 ns (95%) PPS: 200 ns (95%)

The SPS is based on the C/A code and is subject to SA, an accuracy degradation scheme to reduce accuracy to a level within the national security requirements achieved by a combination of orbital parameter dithering. This results in an intentional error regarding the location (ephemeris) of the satellite.

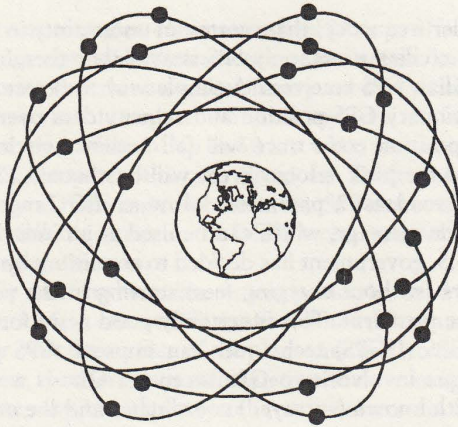


Figure 3.15 Satellite constellation (space segment).

Table 3.3
GPS Characteristics

Item	Characteristic
Satellites	24 satellites broadcast signals autonomously
Orbits	Six planes, at 55-degree inclination, each orbital plane includes four satellites at 20,231-km altitude, with a 12-hr period
Carrier frequencies	L1: 1575.42 MHz; L2: 1227.60 MHz
Digital signals	C/A code (coarse acquisition code): 1.023 MHz P code (precise code): 10.23 MHz Navigation message: 50 bps
Position accuracy	SPS: 100m horizontal (2dRMS) and 140m vertical (95%) PPS: 21m horizontal (2dRMS) and 29m vertical (95%)
Velocity accuracy	SPS: 0.5–2 m/s observed PPS: 0.2 m/s
Time accuracy	SPS: 340 ns (95%) PPS: 200 ns (95%)

The SPS is based on the C/A code and is subject to selective availability (SA). SA is an accuracy degradation scheme to reduce the accuracy available to civilian users to a level within the national security requirements of the United States. It is achieved by a combination of orbital parameter degradation and satellite clock parameter dithering. This results in an intentional error in the navigation message regarding the location (ephemeris) of the satellites and the intentional error in the

timing source (carrier frequency) that creates an uncertainty in the Doppler velocity measurements. For civilian users, only SPS is available, therefore, SA degrades the performance of civilian GPS receivers. A simple way to detect whether SA is on or off is to record stationary GPS position and velocity data over a period of time. If SA is off, the 2D position error trace will fall within a circle with a 25m radius (versus 100 m), and the peak velocity error will be about 0.25 m/s (versus 1 m/s). The satellites also broadcast a parameter known as user range accuracy (URA) as part of the navigation message, which can be used as an indication of the status of SA. Recently, the U.S. government has decided to conditionally phase out SA placed on the civilian users (without charging fees) starting in the year 2000, subject to yearly review by the president. This is certainly good news for GPS civilian users.

Differential GPS (DGPS) techniques can improve GPS performance significantly. The technique involves two GPS receivers. One is a master receiver at a reference station with known (surveyed) coordinates and the other is a receiver at a remote location whose coordinates are to be determined (Figure 3.16). The calculated position solution for the master receiver can be compared to the known coordinates to generate a differential correction for each satellite. In one such commonly used technique, the correction information can then be transmitted to the remote receiver for calculating the corrected local position. This effectively reduces the position error

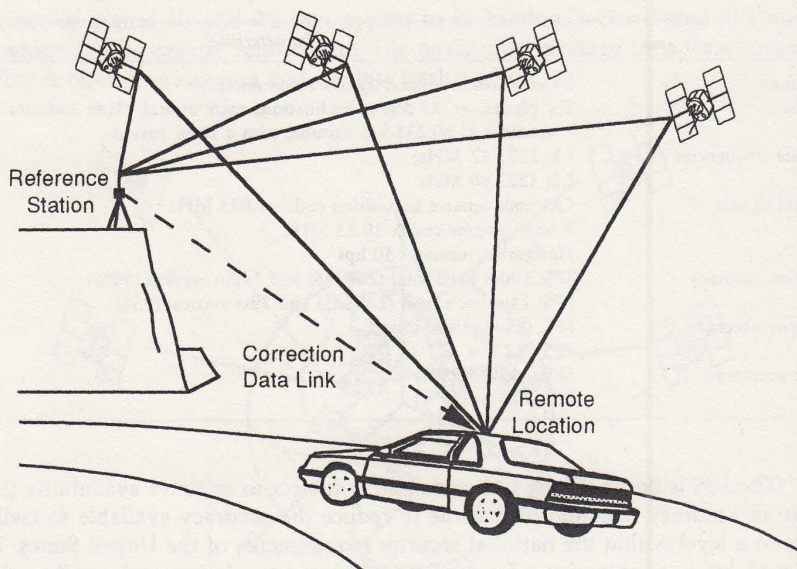


Figure 3.16 Differential GPS.

for commercial systems to under 15m even assuming that the range errors for the remote receiver. It is easy to see that this assumption is only valid when the two receivers are far apart. Accuracy is achieved for receiver separations of up to 1000 km.

Typical DGPS characteristics are shown in Table 3.3. Table 3.3 indicates that DGPS is much better than PPS. A more detailed description of the basic positioning concepts.

We now briefly discuss the measurement of position. As shown in Figure 3.15, a constellation of satellites transmits radio signals to the ground. Radio receivers at or near the surface of the earth receive these signals. These receivers then calculate their position from the signals. This is a one-way broadcast system. Receivers do not send any signals back to the satellites. Since there is an unlimited number of users, much like a television broadcast.

Position measurement is based on the time interval taken for a signal to travel from the satellite to the receiver. The time interval taken for a signal to travel from a known location to reach a receiver is used to obtain the emitter-to-receiver range (distance). This is done by obtaining the time interval from multiple emitters at known locations. The time interval is the clock offset between satellite and receiver. It is impossible to measure the actual range. The clock offset is the constant difference in the satellite and receiver clock. To determine the position, the receiver must know the pseudoranges to the satellites. Pseudoranges are obtained by multiplying the time by the speed of light. The signal propagation time is the time offset required to match the received signal with a replica of the code. This is known as "code correlation." The number of independent observations needed to be solved.

Table 3.3
DGPS Characteristics

Item
Receiver separation
Data link update rate
Position accuracy
Velocity accuracy
Time accuracy

for commercial systems to under 15m even when SA is in effect. DGPS technology assumes that the range errors for the remote receiver are the same as for the master receiver. It is easy to see that this assumption may not be true all the time, especially when the two receivers are far apart. An accuracy level of 5m can commonly be achieved for receiver separations of up to 50 km when low-noise receivers are used.

Typical DGPS characteristics are listed in Table 3.4. Comparison with Table 3.3 indicates that DGPS is much better than SPS and even (surprisingly) better than PPS. A more detailed description of DGPS techniques follows the presentation of the basic positioning concepts.

We now briefly discuss the measurement principles used by the GPS receiver. As shown in Figure 3.15, a constellation of satellites is maintained in Earth orbit. Radio receivers at or near the surface of the planet are used to decode the transmitted satellite signals. These receivers then calculate position, velocity, and time from the signals. This is a one-way broadcast system, so receivers do not need to transmit any signals back to the satellites. Since it is a passive system, GPS can support an unlimited number of users, much like a television broadcast system.

Position measurement is based on the principle of time of arrival (TOA) ranging. The time interval taken for a signal transmitted from an emitter (in our case satellite) at a known location to reach a receiver is multiplied by the speed of the signal to obtain the emitter-to-receiver range (distance). Multiple signals received by a receiver from multiple emitters at known locations are used to determine its location. Because of clock offset between satellite and receiver, propagation delays, and other errors, it is impossible to measure the actual range, so a pseudorange is measured. The clock offset is the constant difference in the time reading between the satellite clock and receiver clock. To determine the position of the receiver, the receiver needs to know the pseudoranges to the satellites being tracked and the positions of those satellites. Pseudoranges are obtained by multiplying the apparent signal propagation time by the speed of light. The signal propagation time is determined by measuring the time offset required to match the received satellite code to an internally generated replica of the code. This is known as "correlation." As in all surveying situations, the number of independent observations required depends on how many unknowns need to be solved.

Table 3.4
DGPS Characteristics

<i>Item</i>	<i>Real-Time Characteristic</i>
Receiver separation	<50 km
Data link update rate	5–10 sec
Position accuracy	15m (2–5m for low-noise receivers)
Velocity accuracy	0.1 m/s
Time accuracy	100 ns

In the 2D case, at least two satellites are required to derive the receiver position. If the positions of the satellites are known and the pseudorange from the receiver to each satellite can be measured, the receiver position is constrained to be one of two points of intersection between circles having radii equal to the pseudoranges (Figure 3.17).

The unknown receiver position (x, y) can be calculated using the following equations. Note that these equations use a 2D Cartesian coordinate system for the reference coordinate frame.

$$p_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2}$$

$$p_2 = \sqrt{(x - x_2)^2 + (y - y_2)^2}$$

where (x_1, y_1) and (x_2, y_2) are the known coordinates of the satellites, and p_1 and p_2 are the measured pseudoranges. After solving these two simultaneous equations, a rough initial estimate (or a third satellite) can be used to determine which of the two solutions should be used.

For the 3D case, three satellites are required (Figure 3.18). Now instead of circles, the pseudoranges yield spheres of position. Clearly, a unique receiver position (x, y, z) of the receiver can be determined by solving the following equations:

$$p_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2}$$

$$p_2 = \sqrt{(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2}$$

$$p_3 = \sqrt{(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2}$$

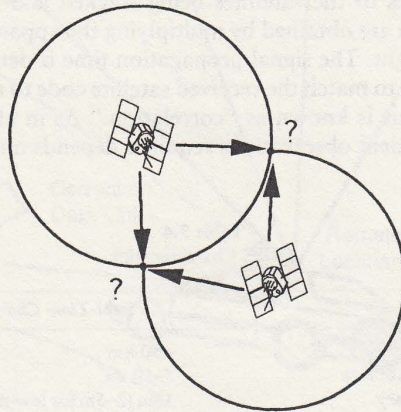


Figure 3.17 Two-pseudorange intersection.

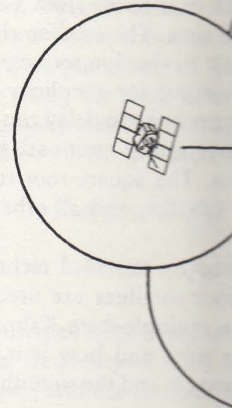


Figure 3.18 Three-pseudorange intersection.

where (x_1, y_1, z_1) , (x_2, y_2, z_2) and (x_3, y_3, z_3) are the known coordinates of the satellites, and p_1 , p_2 and p_3 are the measured pseudoranges.

The receiver clock used to measure time must be synchronized to GPS time. Instead, the clock of the receiver must be determined. This parameter can be determined by design. By design, all of the satellite clocks are synchronized to GPS time. If the receiver clock could be precisely synchronized to GPS time, the measurement would be simple. However, receivers do not use atomic clocks, so instead, inexpensive quartz oscillators introduce a time offset (clock bias) between the receiver clock and the GPS time. This offset must be considered in the computation of the receiver position and clock offset can be derived from the following equations:

$$p_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2}$$

$$p_2 = \sqrt{(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2}$$

$$p_3 = \sqrt{(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2}$$

$$p_4 = \sqrt{(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2}$$

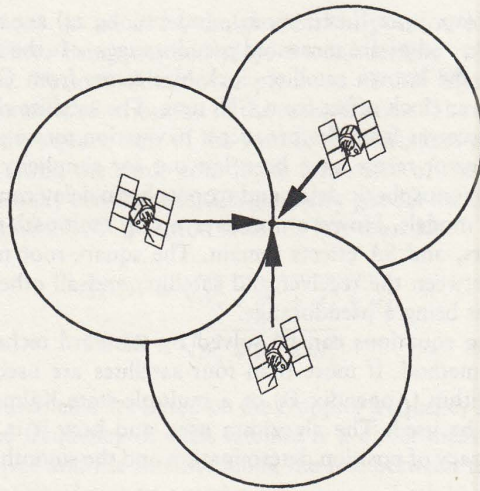


Figure 3.18 Three-pseudorange intersection.

where (x_1, y_1, z_1) , (x_2, y_2, z_2) and (x_3, y_3, z_3) are the known positions of the satellites, and p_1 , p_2 and p_3 are the measured pseudoranges.

The receiver clock used to measure the signal propagation times is not synchronized to GPS time. Instead, the clock offset between receiver time and GPS time must be determined. This parameter can be calculated by adding a fourth satellite. By design, all of the satellite clocks are synchronized using very precise atomic clocks. If the receiver clock could be precisely synchronized with the satellite clocks, time measurement would be simple. However, it is economically impractical for receivers to use atomic clocks, so instead, inexpensive crystal oscillators are used. These clocks introduce a time offset (clock bias) between the receiver and the GPS clocks so we must consider their effects in the computation. The receiver clock bias is the time offset of the receiver, and it is the same for each satellite. Thus both the receiver position and clock offset can be derived from the following equations:

$$p_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2} + c(dt - dT_1)$$

$$p_2 = \sqrt{(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2} + c(dt - dT_2)$$

$$p_3 = \sqrt{(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2} + c(dt - dT_3)$$

$$p_4 = \sqrt{(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2} + c(dt - dT_4)$$

where (x_1, y_1, z_1) , (x_2, y_2, z_2) , (x_3, y_3, z_3) , and (x_4, y_4, z_4) are the known satellite positions, p_1, p_2, p_3 , and p_4 are measured pseudoranges, c is the speed of light, dT_1, dT_2, dT_3, dT_4 are the known satellite clock bias terms from GPS time, and dt is the unknown receiver clock offset from GPS time. The satellite clock bias terms are calculated by the receiver from the broadcast navigation message. In the preceding equations, several error terms have been left out for simplicity. For instance, the range errors due to ionospheric delay and tropospheric delay can both be estimated using atmospheric models. However, receiver noise, multipath propagation error, satellite orbit errors, and SA effects remain. The square-root term represents the geometric range between the receiver and satellite, and all other terms contribute to the measurement being a pseudorange.

The preceding equations can be solved by standard techniques such as the Newton-Raphson method. If more than four satellites are used for a solution, a least squares algorithm (Appendix D) or a multiple-state Kalman filter algorithm (Appendix E) can be used. The algorithm used and how it is used have a large impact on the accuracy of position determination and the smoothness of the receiver trajectory.

Let us use four satellites as an example to explain briefly how to solve these equations and to derive measurements of the satellite geometry with respect to the receiver. After linearization, the above equations can be written in a vector-matrix form as

$$Y = AX$$

where Y is the known vector of measurements and X is the unknown vector. Linear algebra can be used to solve for the unknown vector X . A by-product of the computation is the cofactor matrix

$$P = (A^T A)^{-1}$$

where T stands for the matrix transpose operation and -1 stands for matrix inversion. For a 3D GPS solution, the cofactor matrix is a 4×4 matrix where three components are contributed by the receiver position (x, y, z) and one component by the receiver clock. The matrix can be written as

$$P = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{xt} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} & \sigma_{yt} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 & \sigma_{zt} \\ \sigma_{tx} & \sigma_{ty} & \sigma_{tz} & \sigma_t^2 \end{bmatrix}$$

To use a GPS receiver's output properly, users must understand various potential GPS errors such as the uncertainty in satellite positions and the effects of

propagation conditions. In addition, the respect to the satellites being tracked measurement errors. Various dilution of precision (DOP) is based on the square root of the estimated position along the three coordinates shown below. We usually consider the GDOP (position DOP) is less than 3 or PDOP (position

$$\text{GDOP} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_t^2}$$

$$\text{PDOP} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$$

Velocity measurements are based on the Doppler shift in the frequency of each satellite. The velocity of the receiver and the satellite is a very high velocity relative to a stationary observer. The motion of the satellite and because of the Earth's rotation. The velocity solution can be used in a 3D navigation solution (pseudorange and velocity) with respect to time.

To have a general idea of the complexity of a GPS receiver, Figure 1 depicts a block diagram of a Motorola receiver. The receiver has an application specific integrated circuit (ASIC) for application specific integrated circuit (ASIC) and an erasable programmable ROM. See [21] for more details on the design and some practical receiver examples.

Readers who are interested in evaluating GPS receiver performance in specific applications should consult [22]. In many applications, GPS receiver performance in urban areas is particularly important. Overhead of satellite signals, and high-rise buildings can block satellite signals. Therefore, a receiver with multiple channel tracking (versus multiplexing) is more than one without such features (multiplexing) or all of the available channels among the receiver. For reports under these circumstances, see [24]. Receiver performance is receiver autonomy, receiver detection, isolation, and removal of interference. It indicates that an eight-channel low-cost receiver is needed and performance needed to track satellites.

As mentioned earlier in this section, several errors are associated with the

propagation conditions. In addition, the geometrical position of the receiver with respect to the satellites being tracked plays a considerable role in GPS receiver measurement errors. Various dilution of precision (DOP) measures are used to characterize the geometric contribution to the receiver error. The calculation of these DOPs is based on the square root of the sum of the variances of the errors in the estimated position along the three coordinate axes, and of the estimated time as shown below. We usually consider the GPS geometry good if the GDOP (geometrical DOP) is less than 3 or PDOP (position DOP) is less than 12.

$$\text{GDOP} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_t^2}$$

$$\text{PDOP} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$$

Velocity measurements are based on the Doppler frequency shift principle. The Doppler shift in the frequency of each satellite is a direct measure of the relative velocity of the receiver and the satellite along the line between them. Each satellite has a very high velocity relative to a stationary receiver, both because of the orbital motion of the satellite and because of the motion of the receiver associated with the Earth's rotation. The velocity solution can be calculated by differentiating the four-dimensional navigation solution (pseudorange equations) introduced earlier with respect to time.

To have a general idea of the composition of a GPS receiver, Figure 3.19 depicts a block diagram of a Motorola Oncore GPS receiver, where ASIC stands for application specific integrated circuit and EEPROM stands for electronically erasable programmable ROM. See [21] for more information about the receiver design and some practical receiver examples.

Readers who are interested in evaluation and selection of GPS receivers for specific applications should consult [22] and [23]. For location and navigation applications, GPS receiver performance under heavy foliage and in urban canyon areas is particularly important. Overhead foliage may cause significant attenuation of satellite signals, and high-rise buildings in urban canyons may block and reflect satellite signals. Therefore, a receiver with a position filter, fast reacquisition time, and parallel channel tracking (versus multiplexing channel) capabilities is much better than one without such features (multiplexing is the practice of rapidly sequencing any or all of the available channels among the visible satellites). For performance test reports under these circumstances, see [24,25]. Another feature that enhances GPS receiver performance is receiver autonomous integrity monitoring (RAIM). It provides detection, isolation, and removal of faulty satellite signals. Finally, research indicates that an eight-channel low-cost receiver provides an excellent balance of cost and performance needed to track satellite signals reliably.

As mentioned earlier in this section, when determining the receiver position, several errors are associated with the pseudorange measurement. Table 3.5

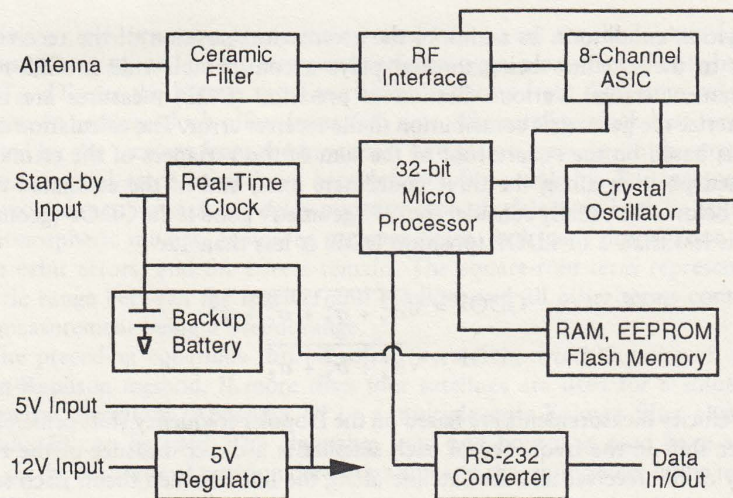


Figure 3.19 Motorola Oncore GPS receiver diagram.

Table 3.5
GPS Receiver Errors

Error	Typical GPS (m)	Worst-Case GPS (m)	Typical DGPS (m)
Satellite clock	2-3	25	<0.5
Satellite orbit	1-2	5	<0.5
SA	30-50	100	<0.5
Ephemeris prediction	3-5	15	<0.5
Ionospheric delay	10-15	100	<0.5
Tropospheric delay	3-5	30	<0.5
Multipath	10	300	10
Receiver noise	5	15	5
Total effect	100 (95%)	300 (99.99%)	15 (95%)

summarizes the typical effects of these errors on the position output by the receiver. From the table, we see that there are other sources of error in addition to SA. The first six errors in Table 3.5 are virtually zero for typical DGPS (receiver separation less than 50 km) because they are highly correlated.

Other techniques can also be used when designing the receiver algorithm to reduce the errors further. For civilian receivers, two methods can be used to calculate receiver position at first lock-on to satellites. One is based on the code phase (1.023 MHz, wavelength $\approx 293,000$ m). The other is based on the carrier phase

(1575.42 MHz, wavelength ≈ 0.19 m) transmitted by the satellite as shown in Figure 3.18. The receiver needs to decide the number of cycles of the code or carrier to correlate with the navigation message. Both methods require the receiver to initially lock on to the signal, which is easily resolved.

Sophisticated algorithms are required to resolve the ambiguity. A combination of these two methods, called code tracking (carrier-smoothed code), is used in the Motorola Oncore DGPS. This method can reduce the error down to 1 to 2 m. One study shows that with carrier phase differential processing and a correction circuit, the accuracy can be made as good as 1 m.

GPS receivers can be used for time synchronization. Recall that the satellites have very expensive atomic clocks with an accuracy (10^{-12}) compared with the best commercial receivers. Furthermore, the clock time-keeping is maintained by other technologies because all satellite clocks are synchronized to GPS time. The time difference between the receiver and the satellites is within 60 ns. Therefore, it is very easy to synchronize time to either GPS time or the receiver's time. Examples of such applications are to synchronize time among three radio-frequency (Sections 9.2.2 and 10.2). Because GPS uses a pulse per second or 1 Hz, these pulses can be used for time synchronization. Using a correction circuit, the receiver can be very precise.

GPS receivers have many different applications. A reference station receiver with few mobile receivers can collect raw data in the field, and the receiver can calculate the distances between the receiver locations in the field. A reference station receiver can be used for a static survey. A reference station receiver can be used for a moving survey. A reference station receiver can be used for a moving survey. A reference station receiver can be used for a moving survey.

DGPS has different implementations. In a static survey, the receiver positions are known, the receiver calculates the range vector from the reference station to the receiver, and then subtracts the known range vector from the

(1575.42 MHz, wavelength ≈ 0.19 m). Both the code and the carrier phases are transmitted by the satellite as shown in Table 3.3. The receiver generates its own code or carrier to correlate with the one received from the satellite in order to decode the navigation message. Both methods have associated range ambiguities because the receiver needs to decide the number of cycles transmitted from the satellite in order for the receiver to initially lock on. For the code phase tracking, the ambiguity is easily resolved.

Sophisticated algorithms are required to resolve the carrier phase tracking ambiguity. A combination of these two observations, as is done in carrier-aided code tracking (carrier-smoothed code) produces better results. For instance, typical Motorola Oncore DGPS can reduce both the multipath and receiver noise errors down to 1 to 2m. One study shows that by postprocessing the raw Oncore output data, DGPS accuracy can be made as good as 2 cm [26]. The postprocessing uses carrier phase differential processing and ambiguity resolution techniques.

GPS receivers can be used for time synchronization and frequency stabilization. Recall that the satellites have very expensive on-board atomic clocks with picosecond accuracy (10^{-12}) compared with the nanosecond accuracy (10^{-9}) of low-cost GPS receivers. Furthermore, the clock time-keeping ability of the satellites is better than other technologies because all satellite clocks are time synchronized by the control segment to GPS time. The time difference between any two GPS receivers is typically within 60 ns. Therefore, it is very economical to use low-cost GPS receivers to synchronize time to either GPS time or universal time coordinated (UTC) time. Examples of such applications are to synchronize time between adjacent cellular telephone cells to smooth the transition of a caller from one cell to another, and to synchronize time among three radio-frequency reference sites to locate a mobile user (Sections 9.2.2 and 10.2). Because GPS receivers can generate steady pulses (one pulse per second or 1 Hz), these pulses can be used for comparison with oscillator frequencies. Using a correction circuit, the frequency of the oscillator can be stabilized very precisely.

GPS receivers have many different applications. In geodetic surveying, a reference station receiver with few mobile receivers at selected surveying locations can collect raw data in the field, and the resulting data can be used to calculate the distances between the receiver locations in the office. This is called a *static postprocessing survey*. A reference station receiver with one mobile receiver can collect data while the mobile receiver is moved to different surveying locations in the field, and the distances are calculated in the office. This is known as a *semi-kinematic* or *rapid static survey*. A reference station receiver or a mobile receiver can broadcast its raw measurements to the other receiver, so that carrier phase differential processing can be done in real time. This method is called a *real-time kinematic (RTK) survey*.

DGPS has different implementations. For instance, if both the satellite positions and the receiver positions are known, the range correction vector can be derived by subtracting the known range vector from the measured range vector. The correction

vector can then be transmitted from the reference station to the mobile receiver. The data transmission rates for the corrections are determined by the rates of variation in the errors. For autonomous vehicle location and navigation, one-way communication of corrections is sufficient. For vehicle tracking, two-way communication may be required if the system needs both the user in the vehicle and a central dispatcher to have corrected positions. VHF (very high frequency) radio, two-way radio, FM subcarrier radio, pager signals, cellular telephone, or satellite-based communications, etc., can be used for this communications link. We discuss RF communication in more detail in Chapter 8.

One cost-effective technique for vehicle tracking is inverse DGPS. In this situation, the vehicle does not have the corrected positions; only the reference station does. This technique requires the vehicle to transmit satellite numbers and ephemeris data for the satellites it is tracking to the reference station along with the vehicle position. The software in the reference station uses these data and other data collected by the reference station using the same satellites to compute a position fix. It then derives the correction vector for that vehicle and applies it. If the system needs to track many vehicles, the reference station needs a very powerful processor or processors because of the computationally intensive nature of the task.

Wide-area DGPS (WADGPS) is another cost-effective method of bringing differential techniques to many users. Recall that DGPS needs to have a receiver separation of less than 50 km to provide very good accuracy. In other words, the main disadvantage of DGPS is the high correlation between the reference-station-to-user distance and the positioning accuracy achieved. To cover a wide area, a substantial increase in the number of reference stations is required. In contrast, WADGPS requires very few reference stations. For example, 10 sparsely located reference stations may be enough to cover the continental United States. Using the GPS receiver data collected by these 10 stations, a national map can be calculated for the entire country. Whenever any local user needs the service, this national map can be used to derive corrections for a virtual reference station in that particular area. Since this virtual reference station is nearly equivalent to a real reference station in the same area, similar position accuracy can be achieved. Research has also been very active in applying similar techniques to determine the exact satellite orbits. If real-time computation can be achieved, this *precise orbits* method should be able to improve the positional accuracy even more. Currently, precise orbits and clocks are available for postmission use only.

In addition to the Global Positioning System developed by the United States, other satellite systems have been developed that could be used for location and navigation, such as GLONASS, which was developed by the former USSR. NAVSAT has been developed by the European Space Agency (ESA), LOCSTAR is being promoted by the French space organization (CNES), and STARFIX has been developed by John E. Chance and Associates of Louisiana for the Gulf of Mexico and continental United States. More information about these systems can be found in [21,27] and

the references cited therein. We discuss in Section 8.3.8; most of the satellite system position determination as a feature. In general, the accuracy of these technologies is worse than that of

3.5 SENSOR FUSION

From the preceding discussion, it is clear that accurate vehicle position information. The system in order to provide the on-vehicle system with information for its location and navigation has been developed to fuse the complementary information from one representation format. The information from multiple sensors during a single period of time or over a period of time. Integrated multisensor systems provide higher levels of accuracy and fault tolerance.

Multisensor integration and fusion provide many advantages. These may include robust operational performance, extended temporal coverage, an increased level of accuracy, enhanced spatial resolution, increased dimensionality, full utilization of sensor capabilities. However, we must keep in mind that although sensor fusion, most fusion methods make certain assumptions. If the assumed sensor model does not adequately represent the sensor, a perfect sensor modeling and estimation result.

3.5.1 Simple Filters

Because of error propagation, many different methods are used. One such method is the digital filter. A low-pass filter can be used to eliminate effects of other magnetic fields on the compass. To eliminate errors due to odometer drift, the odometer data must be integrated to obtain the output from the filter. The filter is a deterministic method of linear system analysis and can be expressed in the form of a transfer function:

$$G(s) = \frac{1}{s\tau + 1}$$

where τ is the time constant of the low-pass filter. τ could be the product of a resistance and a capacitance. We can derive the transfer function for the

the references cited therein. We discuss other satellite communication technologies in Section 8.3.8; most of the satellite systems in operation or in planning also offer position determination as a feature. In general, the positional accuracy provided by these technologies is worse than that of GPS.

3.5 SENSOR FUSION

From the preceding discussion, it is clear that no single sensor can provide completely accurate vehicle position information. Therefore, multisensor integration is required in order to provide the on-vehicle system with complementary, sometimes redundant information for its location and navigation task. Many fusion technologies have been developed to fuse the complementary information from different sources into one representation format. The information to be combined may come from multiple sensors during a single period of time or from a single sensor over an extended period of time. Integrated multisensor systems have the potential to provide high levels of accuracy and fault tolerance.

Multisensor integration and fusion provide a system with additional benefits. These may include robust operational performance, extended spatial coverage, extended temporal coverage, an increased degree of confidence, improved detection performance, enhanced spatial resolution, improved reliability of system operation, increased dimensionality, full utilization of resources, and reduced ambiguity. However, we must keep in mind that although there are many advantages in sensor fusion, most fusion methods make certain assumptions, either explicitly or implicitly. If the assumed sensor model does not adequately describe the data from the real sensor, a perfect sensor modeling and estimation theory may not produce the expected result.

3.5.1 Simple Filters

Because of error propagation, many different error filtration methods are used. One such method is the digital filter. A low-pass filter can be used to reduce the disturbing effects of other magnetic fields on the compass. A high-pass filter can be used to eliminate errors due to odometer drift. The output of these two filters can then be integrated to obtain the output from the positioning module. From the basic deterministic methods of linear system analysis, a low-pass filter can be written in the form of a transfer function:

$$G(s) = \frac{1}{1 + \tau s}$$

where τ is the time constant of the low-pass filter. If a hardware filter is used, τ could be the product of a resistance and a capacitance. By substituting s with $1/s$, we can derive the transfer function for the high-pass filter:

I INTELLIGENT

T TRANSPORTATION

S SYSTEMS



AH Artech House Publishers BOSTON • LONDON