20. Inertial Sensors, GPS, and Odometry

Gregory Dudek, Michael Jenkin

This chapter examines how certain properties of the world can be exploited in order for a robot or other device to develop a model of its own motion or pose (position and orientation) relative to an external frame of reference. Although this is a critical problem for many autonomous robotic systems, the problem of establishing and maintaining an orientation or position estimate of a mobile agent has a long history in terrestrial navigation.

20.1 Odometry

The word *odometry* is a contraction of the Greek words *hodos* meaning *travel* or *journey*, and *metron* meaning measure. Given its importance to a wide variety of applications from civil engineering to military conquest, the basic concepts that underly odometry haven been studied for over 2000 years. Perhaps the earliest reference to odometry appears in the *Ten Books on Architecture* by*Vitruvius*, in which he describes *a useful invention of the greatest ingenuity, transmitted by our predecessors, which enables us, while sitting in a carriage on the road or sailing by sea, to know how many miles of a journey we have accomplished* [20.1]. In the context of autonomous vehicles, odometry usually refers to the use of data from the actuators (wheels, treads, etc.) to estimate the overall motion of the vehicle. The basic concept [20.2] is to develop a mathematical model of how selected motions of the vehicle's wheels, joints, etc. induce motion of the vehicle itself, and then to integrate these specified motions over time in order to develop a model of the pose of the vehicle as a function of time.

The use of odometry information to estimate the pose of the vehicle as a function of time is know as *dead reck-*

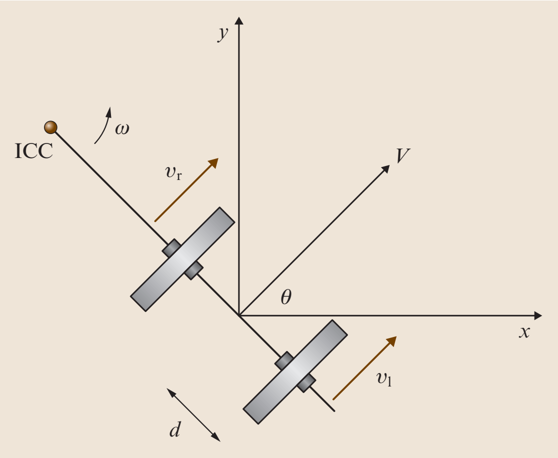

Fig. 20.1 Differential drive kinematics

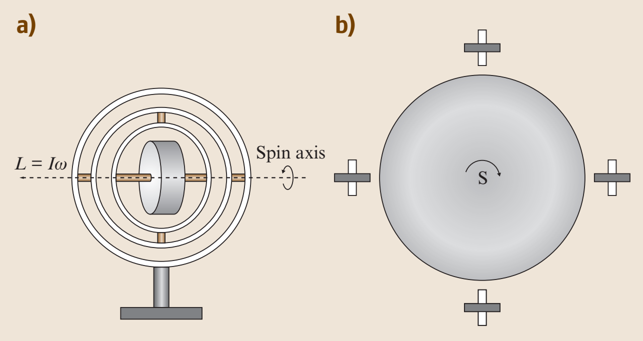

Fig. 20.2a,b Mechanical gyroscope. **(a)** Traditionally gimballed gyroscope. The gimbal provides the gyroscope the freedom to rotate about its axis as the base of the gyroscope is rotated. **(b)** A gyroscope as it is rotated around the planet. The wheel of the gyroscope (*grey*) remains in the same orientation as it revolves with the planet. To an observer on the planet the gyroscope will appear to rotate

oning or *deductive reckoning* and finds wide application in navigation at sea [20.3].

The details of odometry estimation varies by vehicle design. In the context of mobile robots perhaps the simplest vehicle for odometry estimation is the differential drive vehicle (Fig. 20.1). A differential drive vehicle has two driveable wheels which are independently controllable and which are mounted along a common axis. Assuming that the location of the wheels are fixed on the vehicle, then for the wheels to remain in constant contact with the ground, the two wheels must describe arcs on the plane such that the vehicle rotates around a point (known as the ICC – instantaneous center of curvature) that lies on the wheels' common axis (Fig. 20.1). If the ground contact speeds of the left and right wheels are v_l and v_r respectively, and the wheels are separated by a distance $2d$, then

$$\omega(R+d) = v_\mathrm{l}$$
$$\omega(R-d) = v_\mathrm{r}\,.$$

We can rearrange these two equations to solve for ω the rate of rotation about the ICC and R the distance from the center of the robot to the ICC

$$\omega = \frac{(v_\mathrm{l} - r_\mathrm{r})}{(2d)}\,,$$
$$R = d\frac{(v_\mathrm{r} + v_\mathrm{l})}{(v_\mathrm{l} - v_\mathrm{r})}\,.$$

The instantaneous velocity of the point midway between the robot's wheels is given by $V = \omega R$.

Now as v_l and v_r are functions of time we can generate a set of equations of motion for the differential drive robot. Using the point midway between the wheels as the origin of the robot, and writing θ as the orientation of the robot with respect to the x-axis of a global Cartesian coordinate system, one obtains

$$x(t) = \int V(t)\cos(\theta(t))\,\mathrm{d}t\,,$$
$$y(t) = \int V(t)\sin(\theta(t))\,\mathrm{d}t\,,$$
$$\theta(t) = \int \omega(t)\,\mathrm{d}t\,.$$

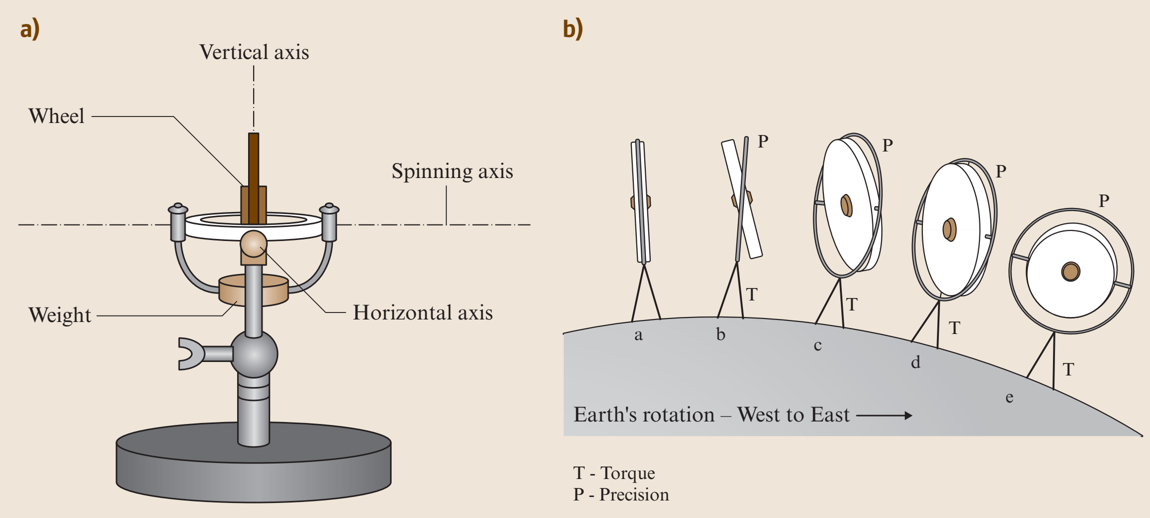

Fig. 20.3a,b Simple gyrocompass. **(a)** Pendulus gyro. **(b)** Precessional motion

This is the solution for the odometry of a differential drive vehicle on the plane. Given the control inputs (v_l and v_r) and some initial state estimate, we can estimate the state of an idealized robot using this motion model at any time t.

Given such a model and complete knowledge of the control inputs, we should, in principle, be able to estimate a robot's pose at any time. In a perfect world this would be all that is necessary to estimate accurately the robot's pose at any time in the future. Unfortunately errors in the modeling (incorrect estimations of wheel size, vehicle size), uncertainty about the control inputs, realities of the motor controller (errors between commanded wheel rotation and true rotation), errors in the physical modeling of the robot (wheel compaction, ground compaction, wheel slippage, nonzero tire width), etc., introduce an error between the dead reckoning estimate of the vehicle motion and its true motion. The problem of correcting for this error is the problem of *pose maintenance* for the vehicle, and requires the integration of the dead reckoning estimate with estimates obtained from other sensor systems.

Other chapters in this handbook (e.g., Chaps. 21–24) examine sensors that rely on external events, visual and otherwise, that can provide information as to the robot's pose or changes in it's pose. Here we consider sensors that transduce physical properties of matter under the influence of external forces and properties of matter and the use of a global position system (GPS).

20.2 Gyroscopic Systems

The goal of gyroscopic systems is to measure changes in vehicle orientation by taking advantage of physical laws that produce predictable effects under rotation. A rotating frame is not an inertial frame, and thus many physical systems will appear to behave in an apparently *non-Newtonian* manner. By measuring these deviations from what would be expected in a Newtonian frame the underlying self-rotation can be extracted.

20.2.1 Mechanical Systems

Mechanical gyroscopes and gyrocompasses have a long history in navigation, *Bohnenberger* is generally credited with the first recorded construction of a gyroscope [20.4], and in 1851 Léon Foucault recognized the gyroscope as an inertial frame. The gyrocompass was patented in 1885 by Martinus Gerardus ven den Bos. In 1903 Herman Anschütz-Kaempfe constructed a working gyrocompass and obtained a patent on the design. In 1908 Elmer Sperry patented a gyrocompass in the US and then attempted to sell this device to the German Navy. A patent battle followed, and Albert Einstein testified in the case. (See [20.5–8] for more details on the history of the gyrocompass and its inventors.)

Gyroscopes and gyrocompasses rely on the principle of the conservation of angular momentum [20.9]. Angular momentum is the tendency of a rotating object to keep rotating at the same angular speed about the same axis of rotation in the absence of an external torque. The angular momentum $\boldsymbol{L}$ of an object with moment of inertia $\boldsymbol{I}$ rotating at angular speed ω is given by

$$\boldsymbol{L} = \boldsymbol{I} \times \boldsymbol{\omega} .$$

Consider a rapidly spinning wheel mounted on a shaft so that it is free to change its axis of rotation (Fig. 20.2a). Assuming no friction due to air resistance or the bearings, the rotor axis will remain constant regardless of the motion of the external cage. This constancy of orientation can be exploited to maintain a bearing independently of the motion of the vehicle, although it is not usually desirable to use the principle of conservation of angular momentum via a gyroscope directly. To see this, suppose that a gyroscope is set on the equator, with its spinning axis aligned along the equator (Fig. 20.2b). As the Earth spins, the gyroscope will maintain a constant axis of orientation and thus to an Earth-fixed observer will appear to rotate, returning to its original orientation every 24 h. Similarly, if the gyroscope were to be positioned on the equator such that its spinning axis was parallel to the axis of rotation of the earth, the gyroscope's axis of rotation would remain stationary and would appear to remain stationary to an Earth-fixed observer as the planet rotates.

Although this global motion limits the mechanical gyroscope's ability to sense absolute orientation directly, gyroscopes can be used to measure local changes in orientation, and thus are well suited to vehicular robotic applications. Rate gyros (RGs) measure a vehicle's rotation rate (its *angular rate* of rotation). This is the fundamental measurement that is the basis of all gyroscopic systems. Rate-integrating gyros (RIGs) use embedded processing to internally integrate the angular rotation rate to produce an estimate of the absolute angular displacement of the vehicle.

In order to exploit a gyroscope for navigation with respect to an Earth-stable frame, it is desirable for the rotational axis of the shaft to remain fixed within the Earth frame, rather than remaining fixed with respect to an external frame. A gyrocompass obtains this by relying on precession. When a torque is applied to change the axis of rotation of a spinning object, conservation of angular momentum will cause a change in the spin direction that is perpendicular to the angular momentum and the direction in which the torque is applied. This is the effect that causes gyroscopes suspended at one end to *spin* around the end from which they have been suspended. Consider the *pendulus gyro* sketched in Fig. 20.3a. This is a standard gyroscope with a weight suspended below the rotational axis. As before, imagine this pendulus gyrocompass set spinning on the equator with the axis of rotation aligned with the axis of rotation of the planet, and with the weight hanging directly down. As the planet spins, the gyroscope's axis of rotation remains stationary and would appear to remain stationary as the planet rotates. Now suppose that instead of being aligned with the spin axis of the planet the spin axis is aligned with the equator. As the planet spins, the spin axis is drawn out of the plane as it remains aligned with the original spin axis. As it becomes drawn out of the plane, the mass hanging below the gyrocompass is raised up and generates a torque down due to gravity. The direction perpendicular to the spin axis and the torque rotates the spin axis away from the equator known and towards the true pole. This process is sketched in Fig. 20.3b.

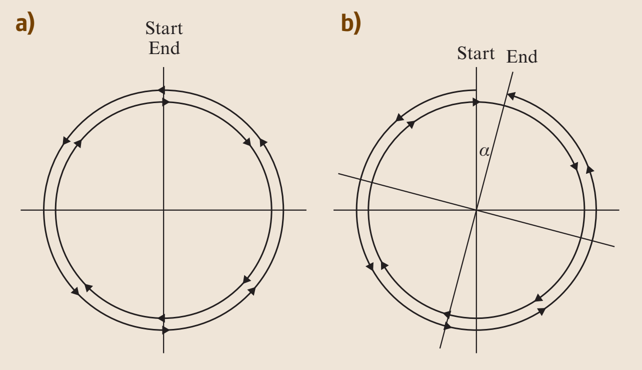

Fig. 20.4a,b Circular light path. **(a)** Stationary path. **(b)** Moving path

Unfortunately the pendulus gyro is not an ideal device for navigation. Although its rotation axis will align with the rotation axis of the planet, it does not converge to this value directly but oscillates about it. The solution to this damping problem is to use oil reservoirs, rather than a solid weight, as the counterbalance, and to restrict the motion of the oil in the reservoir [20.10].

The gyrocompass finds true north by controlling the precession of a gyroscope. In practice, the performance of a mechanical gyrocompass is impacted by external forces acting on the compass which also contribute to the precession of the gyroscope. This includes forces generated by motion of the device containing the gyrocompass, as well as any external forces acting on the vehicle. Another issue for mechanical gyrocompasses is that in latitudes away from the equator the stable position of the gyrocompass is not horizontal, and accurate estimates of true north at such latitudes requires corrections to be applied to raw gyrocompass values. Finally, mechanical gyrocompasses require an external force to be applied to maintain the spin of the gyroscope. This process introduces unwanted forces into the system which can further corrupt the measurement process.

Given the complexity, cost, size, the delicate nature of gyrocompasses, and the availability of less expensive and more reliable technologies, mechanical gyrocompasses have given way to optical- and microelectromechanical systems (MEMS)-based systems.

20.2.2 Optical Systems

Optical gyroscopes rely on the *Sagnac effect* rather than rotational inertia in order to measure (relative) heading. The mechanism is based on the behavior of an optical standing wave in a rotating frame. Historically this was first produced using lasers and an arrangement of mirrors, but it is now typically obtained using fibreoptic technology. The Sagnac effect is named after its discoverer Georges *Sagnac* [20.11, 12]. The underlying concept can be traced back even earlier to the work of *Harress* [20.13], and perhaps finds its most famous application in terms of the measurement of the rotation of the Earth [20.14].

Ignore relativistic effects and consider the circular light path shown in Fig. 20.4a. If two light pulses are sent in opposite directions around a stationary path of perimeter $D = 2\pi R$ they will travel the same distance at the same speed. They will arrive at the starting point simultaneously, taking time $t = D/c$ (where c is the speed of light in the medium). Now let us suppose that instead of being stationary, this circular light path rotates clockwise about its center at rotational speed ω (Fig. 20.4b). The light traveling clockwise around the path must go *farther* to reach the starting point, while light traveling counterclockwise around the path goes a *shorter* distance. The clockwise path has distance $D_c = 2\pi R + \omega R t_c$, where t_c is the time taken in the clockwise direction, while the counterclockwise path has distance $D_a = 2\pi R - \omega R t_a$, where t_a is the time taken in the counterclockwise direction. But $D_c = ct_c$ and $D_a = ct_a$, so $t_c = 2\pi R/(c - \omega R)$ and $t_a = 2\pi R(c + \omega R)$. The time difference $\Delta t = t_c - t_a$ is given by

$$\Delta t = 2\pi R \left(\frac{1}{c - \omega R} - \frac{1}{c + \omega R} \right) .$$

By measuring Δt, the rotational speed can be computed. Note that, although the above derivation assumes classical mechanics and ignores relativistic effects, the derivation also applies when relativistic speeds are taken into account [20.15]. See [20.16] for an in-depth review of the Sagnac effect and ring lasers.

In optical gyroscopes lasers are typically used as the light source. Optical gyroscopes either employ straight-line light paths with mirror surfaces or prisms at the edges to direct the light beam (a ring laser gyroscope – RLG), or a polarization maintaining glass-fiber loop (fiber optic gyro – FOG). The glass fiber may actually loop multiple times, thus extending the effective length of the light path. The time delay between the clockwise and counterclockwise directions is detected by examining the phase interference between the clockwise and counterclockwise light signals. Multiple optical gyroscopes with nonparallel axes can be ganged together in order to measure three-dimensional (3-D) rotations.

Various techniques can be used to measure the time difference between the two paths, including examining the Doppler (frequency) shift of the laser light due to the motion of the gyro and an examination of the beat frequency of the interference pattern between the clockwise and counterclockwise paths [20.17]. Ring interferometers typically consist of many windings of fiber-optic lines that conduct light of a fixed frequency in opposite directions around the loop and measure the phase difference. A ring laser consists of a laser cavity in the shape of a ring. Light circulates in both directions around this cavity, producing two standing waves with the same number of nodes in both directions. Since the optical path lengths are different in the two directions, the resonant frequencies differ. The difference between these two frequencies is measured. An unfortunate side-effect of the ring-laser approach is that the two signals will lock in to each other for small rotations and it is typically necessary to physically rotate the device in a controlled manner in order to ensure that this lock-in effect can be avoided.

20.2.3 MEMS

Almost all micro-electromechanical systems (MEMS) gyroscopes are based on vibrating mechanical elements to sense rotation. Vibratory gyroscopes rely on the transfer of energy between vibratory modes based on Coriolis acceleration. Coriolis acceleration is the apparent acceleration that arises in a rotating frame of references. Suppose that an object moves along a straight line in a rotating frame of reference. To an outside observer in an inertial frame the object's path is curved – thus there must be some force acting on the object to maintain the straight line motion as viewed by the rotating observer. An object moving in a straight line with local velocity $\boldsymbol{v}$ in a frame rotating at rate $\boldsymbol{\Omega}$ relative to an inertial frame will experience a Coriolis acceleration given by

$$\boldsymbol{a} = 2\boldsymbol{v} \times \boldsymbol{\Omega} .$$

Transducing acceleration in a MEMS gyroscope amounts to inducing some local linear velocity and measuring the resultant Coriolis forces.

Early MEMS gyroscopes utilized vibrating quartz crystals to generate the necessary linear motion. More recent designs have replaced the vibrating quartz crystals with silicon-based vibrators. Various MEMS structures have been developed including those described below.

Tuning-Fork Gyroscopes

Tuning-fork gyroscopes use a tuning-fork-like structure (Fig. 20.5) as the underlying mechanism. As the tuning forks vibrate within a rotating frame, Coriolis forces cause the tines of the fork to vibrate out of the plane of the fork, which is measured. This is the effect used by the InertiaCube sensor [20.18].

Vibrating Wheel Gyroscopes

Vibrating wheel gyroscopes use a wheel that oscillates about its rotational axis. External rotation of the frame causes the wheel to tilt, which can be sensed.

Wine-Glass Resonator Gyroscopes

Wine-glass resonator gyroscopes use the effect of Coriolis forces on the position of nodal points on a resonating structure to estimate the external rotation.

As MEMS gyroscopes have no rotating parts, have low-power consumption requirements, and are very

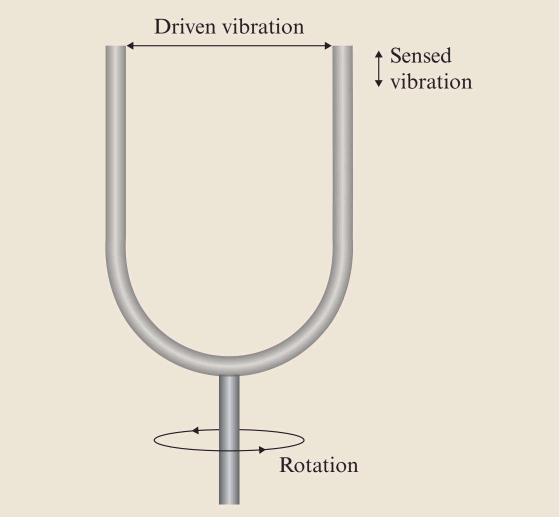

Fig. 20.5 MEMS gyroscope: principle of operation

small, MEMS gyros are quickly replacing mechanical and optical gyroscope sensors in robotic applications.

20.2.4 Performance

Inertial measurement units can be evaluated with respect to various factors that determine performance, several of which are enumerated below.

1. Bias repeatability. This is the maximum deviation of the gyroscope under fixed inertial operation conditions with constant temperature, i.e., the drift of the reading under ideal conditions. This is measured over different time scales and leads to *short-term* and *long-term* bias repeatability.
2. Angle random walk. This measures the noise in the angular rate data coming from the gyro.
3. Scale factor ratio. This parameter is not specific to inertial measurement units (IMUs) or gyros and is a general measurement of signal amplitude. It measures the ratio of the output analogue voltage to the sensor parameter of interest. For a gyroscope this is typically measured in mV/(deg/sec) whereas for an accelerometer it is typically measured in $\mathrm{mV/(m/s^2)}$.

20.2.5 Summary

With the exception of gyrocompasses, gyroscopes measure relative rotational motion around a single axis. They accomplish this measurement by exploiting physical properties of rotating frames of reference. Earlier technologies based on mechanical gyroscopes have given way to optical- and MEMS-based devices but the underlying principle remains unchanged: that rotating frames of reference show specific physical properties that can be measured to estimate the relative rotation.

A problem common to all gyroscopes is that of drift. Each of the relative motion measurements is corrupted by an error process, and these errors accumulate over time. This, coupled with specific measurement errors associated with the individual gyroscope technologies, means that unless the error is corrected through reference to some alternate (external) measurement, the drift will eventually exceed the required accuracy of the measurement.

As individual gyros only measure rotation about a single axis, it is common to gang multiple gyros together with orthogonal axes of sensitivity in order to measure 3-D rotations. These collections of gyros are often integrated with other sensors (compasses, accelerometers, etc.) in order to construct *inertial measurement units* (or IMUs). This is considered in Sect. 20.4.

20.3 Accelerometers

Just as gyroscopes can be used to measure changes in orientation of a robot, other inertial sensors – known as *accelerometers* – can be used to measure external forces acting on the vehicle. One important factor concerning accelerometers is that they are sensitive to all external forces acting upon them – including gravity. Accelerometers use one of a number of different mechanisms to transduce external forces into a computer-readable signal.

Mechanical Accelerometer

A mechanical accelerometer (Fig. 20.6a) is essentially a spring–mass–damper system with some mechanism for external monitoring. When some force is applied (e.g., gravity), the force acts on the mass and displaces the spring. Assuming an ideal spring with a force proportional to its displacement, the external forces balance the internal ones

$$\begin{aligned} F_{\text{applied}} &= F_{\text{inertial}} + F_{\text{damping}} + F_{\text{spring}} \\ &= m\ddot{x} + c\dot{x} + kx \,, \end{aligned}$$

where c is the damping coefficient. This equation can be solved to show that, depending on the size of the damping coefficient relative to the expected external force and the mass, the system can be made to reach a stable final value in a reasonably short period of time whenever a static force is presented. This need to pre-estimate the expected force and the resulting (potentially long) time for the system to converge on a final measurement coupled with nonideal performance of the spring limits the applicability of mechanical accelerometers. Another issue with mechanical accelerometers is that they are particularly sensitive to vibration.

Piezoelectric Accelerometer

Rather than relying on a direct mechanical measurement of external forces, piezoelectric accelerometers are based on a property exhibited by certain crystals, across which a voltage is generated when they are stressed. A small mass can be positioned so that it is only supported by the crystal, and as forces cause the mass to act upon the crystal this induces a voltage that can be measured (Fig. 20.6(b)).

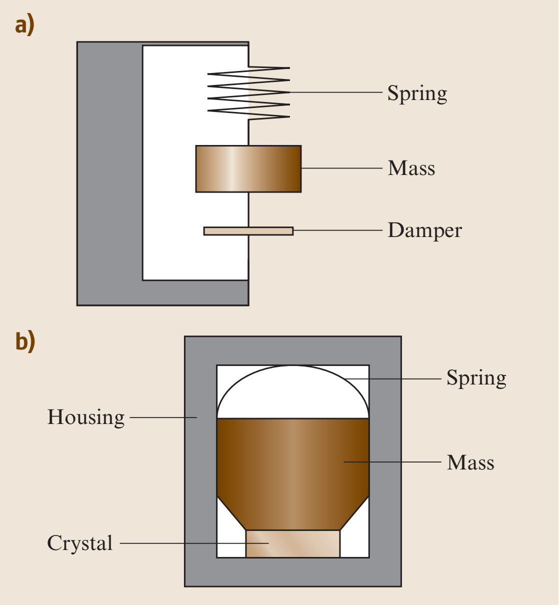

Fig. 20.6a,b Accelerometers. **(a)** Mechanical accelerometer. **(b)** Piezoelectric accelerometer

20.4 IMU Packages

An inertial measurement unit (IMU) is a device that utilizes measurement systems such as gyroscopes and accelerometers to estimate the relative position, velocity, and acceleration of a vehicle in motion. The resulting navigation system is known as an inertial navigation system or INS. First demonstrated in 1949 by C. S. Draper, IMUs have become a common navigational component of aircraft and ships. Historically an IMU is self-contained and provides this estimate without reference to external references, however the definition has become less precise in recent years and now it is common to apply the term IMU also to systems that do include such external references.

IMUs come in two basic flavors, gimbaled systems and strap-down systems. As their name suggests, gimbaled IMUs are mounted within complex gimbal structures in order to provide a stable platform from which measurements can be made. Gyroscopes are used to ensure that the gimbal remains aligned with the initial reference frame at power up. The orientation of the gimbaled platform relative to the vehicle is used to map measurements taken within the IMU to the reference frame of the vehicle. Strap-down IMUs, on the other hand, have the IMU rigidly connected to the vehicle (strapped down), so no such transformation is required. In either case estimating the motion relative to the initial frame requires integrating information from the sensors within the IMU (accelerometers, gyroscopes, etc.) in real time. This was a significant computational effort in the early days of IMUs, and thus historically (prior to the 1970s) the gimbaled IMU was more common. Given the low cost of such computation today, and the costs associated with manufacturing and operating gimbaled IMUs, strap-down IMUs are much more common today [20.19].

A true IMU maintains a 6-degree-of-freedom (DOF) estimate of the pose of the vehicle: position (x, y, z) and orientation (roll, pitch, yaw). IMU-like systems that (for example) only maintain ongoing estimates of orientation

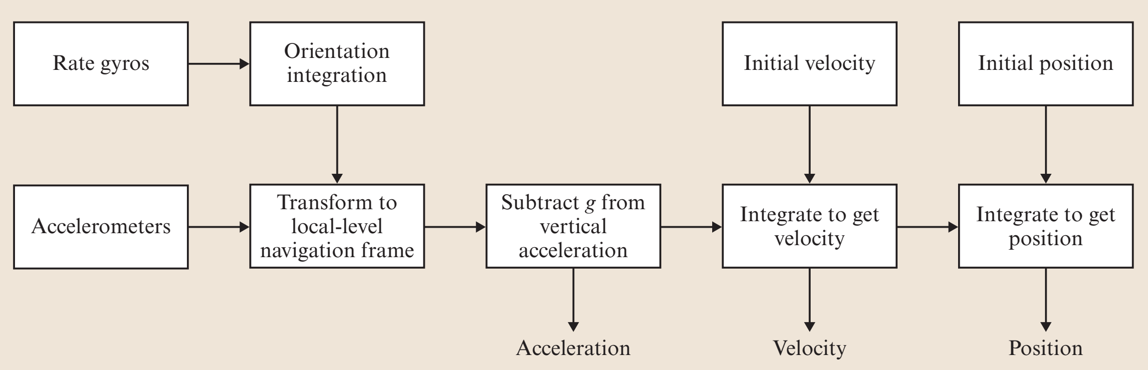

Fig. 20.7 IMU block diagram

are known as *attitude and heading reference systems*, and operate in a similar fashion as IMUs but maintain a less full representation of the vehicle's state. In addition to maintaining a 6-DOF pose of the vehicle, commercial IMUs also typically maintain estimates of velocity and acceleration.

The basic computational task of an IMU is shown in Fig. 20.7. This IMU uses three orthogonal accelerometers and three orthogonal gyroscopes. The gyroscope data $\boldsymbol{\omega}$ is integrated to maintain an ongoing estimate of vehicle orientation $\boldsymbol{\theta}$. At the same time, three accelerometers are used to estimate the instantaneous vehicle acceleration $\boldsymbol{a}$. This data is then transformed via the current estimate of the vehicle orientation relative to gravity, so that the gravity vector can be estimated and extracted from the measurement. The resulting acceleration is then integrated to obtain vehicle velocity $\boldsymbol{v}$ and then integrated again to obtain the position $\boldsymbol{r}$.

IMUs are extremely sensitive to measurement errors in the underlying gyroscopes and accelerometers. Drift in the gyroscopes leads to misestimates of the vehicle orientation relative to gravity, resulting in incorrect cancellation of the gravity vector. As the accelerometer data is integrated twice, any residual gravity vector will result in a quadratic error in position [20.18]. As it is never possible to eliminate the gravity vector completely, and this and any other error is integrated over time, drift is a fundamental issue for any IMU. Given a sufficiently long period of operation all IMUs eventually drift and reference to some external measurement is required to correct this. For many field robots GPS has become an effective source for these external corrections.

20.5 GPS

The global positioning system (GPS) is the single most commonly used mechanism for location estimation. It provides a three-dimensional position estimate in absolute coordinates as well as current time and date and is available anywhere on the Earth's surface. Standard GPS provides a position estimate in the horizontal plane to within about 20 m. It was original developed for military applications but has become widely adopted in civilian applications, including automobile navigation systems, recreational orienteering, and inventory tracking for transportation companies.

20.5.1 Overview

The system is based on received radio signals transmitted by an ensemble of satellites orbiting the Earth. By comparing the time delays from the different satellite signals, a position fix can be computed. The most widely accepted GPS system is based on the *NAVSTAR* satellite system deployed and maintained by the United States, specifically by the Air Force Space Command. As a US military service, the US government reserves the right to terminate or modify its availability at their discretion. A similar system named *Globalnaya Navigatsionnaya Sputnikovaya Sistema* (*GLONASS*) is operated by the Russian government, but at the time of writing it is not available for robotic applications. Another alternative system is being deployed by the European Union, named *Galileo*, which the explicit expectation that it will not be under military control. It is expected to offer two different classes of service: an open service and an encrypted higher-quality commercial service. Other GPS systems such as the Chinese *Beidou* or the Japanese *QZSS* are outside the scope of this chapter.

In general the term GPS always refers to the NAVSTAR system. Historically NAVSTAR provided two different services: the precise positioning system (PPS), reserved primarily for military users, and the standard position system (SPS) with lower accuracy. The difference between these was referred to as selective availability (SA). This difference in accuracy was artificially induced via pseudorandom noise in the SPS signal for strategic reasons, and was eliminated in 2001. Although this accuracy distinction could, in principle, be reinstated this seems unlikely due to the widespread commercial dependence on GPS that has arisen in the last decade. Even now, the GPS network includes supplementary data in encrypted precision P(Y) code which is not available to civilian users.

The GPS satellite network is based on a base constellation of 24 orbiting satellites along with up to six supplementary additional satellites that are also operational. These satellites are in almost-circular medium Earth orbit. As opposed to being geostationary the orbits are semisynchronous, meaning that their position relative to a ground observer is constantly changing, and that their orbital period is exactly half a sidereal day. The orbits are selected so that from almost any point of the Earth's surface there will always be four or more satellites directly visible – a criterion for obtaining a GPS position estimate. The satellites are organized into six orbital planes with four satellites in each. The system is designed such that averaged over the entire Earth's surface and over a 24 h internal, the satellites should allow 99.9% coverage at the worst-covered location in any 24 h interval, and the signal should be at least 83.92% available at the worst place on Earth on the worst day over a 30-day measurement interval. Note that this criterion for availability takes into account transmitted operational factors above and beyond simple coverage. Of course, this criterion ignores the reality of topographic features like mountains, as well as other objects such as buildings that can obstruct the line of sight.

Each satellite repeatedly broadcasts a data packet known as the coarse-acquisition (C/A) code, which is received by the GPS receiver on the L1 channel at 1575.42 MHz. The simple principle is that, if the receiver knows the absolute positions of the observed satellites, the receiver position can be directly determined. If the signal propagation time for the radio signals were known, the receiver position could be computed directly via trilateration (Fig. 20.8). This implies that an absolute timing reference is present on the receiver, which would be prohibitively costly. Instead, only the satellites have highly accurate atomic clocks (accurate to approximately 1 s in 300 000 years). The receiver computes the difference in signal propagation times between the different satellites, and uses this to compute a range estimate referred to as a *pseudorange* (to explicitly indicate that it is corrupted by several sources of measurement noise). The specific geometric problem is referred to as *multilateration* or *hyperbolic positioning* and the solution is computed using a sophistical Kalman filter within the GPS receiver. To avoid retaining an ephemeris (pose) table for the satellites and a very accurate clock in the receiver, each satellite broadcasts its own position and an accurate time signal as part of the data packet that it transmits.

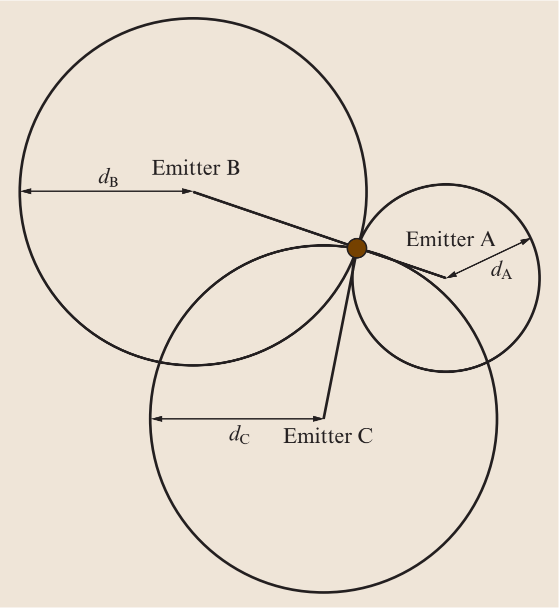

Fig. 20.8 GPS trilateration on the plane. Suppose that one receives signals from three transmitters (A, B, and C) with known locations. Knowledge of the signal delay from one emitter (say A) localizes the receiver to lie on a circle of known diameter (d_A) whose center is the emitter. The constraints from two emitters intersect at two points (maximum). A third emitter is required to disambiguate these two solutions. In three dimensions, the signal propagation constraint from a single emitter constrains the receiver to lie on a sphere. The intersection of the constraints from two emitters constrains the receiver to lie on a circle. The intersection of the constraints from three emitters constraints the receiver to one of two points

GPS satellites broadcast at several different frequencies known as L1 through L5; only L1 (1575.42 MHz) and L2 (1227.6 MHz) are used by civilian GPS receivers. The standard service offered by NAVSTAR and the performance criteria for it are determined by the L1 signal, which contains two unencrypted components: the acquisition message (coarse-acquisition message C/A) and a navigation data message. It is also possible to use the encrypted L2 signal as well, even without the secret decryption keys, to provide augmented error correction (by observing the relative effects of ionospheric distortion as a function of frequency). The restricted-access signal broadcast on both the L1 and L2 channels is the P-code (as well as a fairly recent M-code) which is known as the Y-code or P(Y) or P/Y code once it is encrypted. Both the C/A and P(Y) codes include the navigation message stream that specifies clock bias data, orbital information, ionospherical propagation corrections factors, ephemeris data, status information on all the satellites, universal time code, and other information. The satellite performance is coordinated by the master control station located at the Schriever Air Force Base near Colorado Springs, Colorado, USA and is connected to a global network of five additional monitoring stations (Cape Canaveral USA, Ascension Island, Kwajalein Atoll in the Marshall Islands, Diego Garcia Atoll, and Hawaii) which are used produce the measurements that are uplinked to generate the navigation message stream. Finally, it should be noted that an additional signal is now available on the L2 frequency band. This L2C signal on satellites designated *block IIR-M* promises to provide much improved receiver sensitivity so that position fixes can be obtained in environments, such as in forests, where they are currently not readily available.

20.5.2 Performance Factors

GPS performance depends on several factors: satellite transmission accuracy, environmental conditions, interactions with ground-based obstacles, and receiver properties.

In the context of robotics, factors that affect the performance of the satellites themselves and the atmospheric conditions are essentially uncontrollable. Nevertheless, it should be noted that these can be sources of error and that the GPS signal itself may not always be reliable. A *service failure* is defined as a set of circumstances where the positioning service exhibits atypical error behavior (i. e., incorrect signals). Such failures are classified as either minor and major failures. Minor failures are those that have limited impact on a receiver and lead to ranging errors of under 150 m. Major failures are those that lead either to larger errors or data processing overloads in the receiver. If a single satellite experiences an error that leads to a major failure, then within a 6 h period, approximately 63% of the Earth's surface will have the satellite in view at some point.

The controllable factors in using GPS for accurate localization are

1. it requires an unobstructed line of sight to the satellites,
2. it depends on atmospheric conditions, and
3. it depends on the ability to receive (weak) radiofrequency communications.

There is a potential for wildly incorrect estimates. Generally satellites that are directly overhead provide better signals than those near the horizon. In addition, since the basis of GPS position is differential signal analysis, it is best if the satellites used in the GPS computation are widely spaced in the sky.

GPS signals are in the microwave band and, as such, they can pass through plastic and glass, but are absorbed by water (wood, heavy foliage) and are reflected by many materials. As a consequence, GPS is unreliable in heavy forest, deep canyons, inside automobiles and boats, in heavy snowfall or between tall buildings. In some cases, partial obstruction of the sky may not prevent a position estimate from being computed. Assuming the minimum number of satellites operating at any time is 24, then on average across the Earth's surface eight satellites are in view so that even partial occlusion of the sky can often be tolerated. On the other hand, partial occlusion can lead to reduced accuracy since the selection of available satellites used for computing position becomes limited and optimal accuracy is obtained by using as many satellites as possible (weighting them appropriately in the internal Kalman filter).

Secondary factors that differentiate different GPS receivers are the rate at which the signals are collected, the receiver sensitivity, the number of satellites used in the final computation, the number of factors taken into account in the estimator, and the exploitation of supplementary positioning schemes such as the wide-area augmentation system (WAAS) (see later). A major factor in determining the rate at which estimates can be produced is the number of independent receiver elements in the GPS system. Sequential single-channel receivers are simpler and thus more economical (and potentially smaller), but they must lock sequentially onto each satellite being used. Parallel multichannel receivers can lock onto to more than one satellite at once, and are generally faster and more costly; some degree of parallelism is the norm in good-quality consumer devices.

GPS computations are based on an estimation of the dilution of precision (DOP) and specifically for the dilution of precision of the positional parts of the system, i. e., positional dilution of precision (PDOP). These correspond to partial derivatives of the error with respect to position and allow the most accurate ensemble of visible satellites to be determined at any time. The standard implementation for GPS systems specifies that PDOP values be recomputed every 5 min.

The minimum performance parameters for GPS receivers are based upon transforming instantaneous range residuals to a user position estimate using a linearized position solution from a stationary, surveyed location. Most GPS receivers use additional techniques such as range residual smoothing, velocity aiding, Kalman filtering, or multiple satellite (*all-in-view satellite*) solutions. That said, formal performance for the system is measured with respect to the minimum. The GPS position estimation algorithm is summarized as follows

1. Select the best four satellites based upon the minimum error measured in terms of PDOP.
2. Update every five minutes, or whenever a satellite being used in the solution sets.
3. Measure the pseudorange to each satellite. Each of the four measurements must have a reception time tag within 0.5 s of the solution time. The reception time tag is based upon measurement system time, and the transmission time tag is based upon satellite time.
4. Determine the ephemeris for each of the satellites being used, and compute the Earth-centered, Earth-fixed (ECEF) coordinates for each. Correct for the Earth's rotation and thus compute an estimated pseudorange measurement that should be obtained for each satellite.
5. Compute the *range residuals* as the differences between the actual and observed measurements.
6. Estimate the matrix G that determines the overall system solution, know as the *position solution geometry matrix*. The matrix can be described in terms of a collection of row vectors, one for each of the satellites being used, each row being made up of the x, y, z and time coordinate direction cosines for the vector between the user and the satellite (with respect to a fixed reference frame for the planet called the World Geodetic System, WGS84).
7. Compute the user's position.

The standard implementation of GPS is based upon a position fix rate of once per second, although faster and slower rates are possible. Under typical operating conditions and without specialized enhancements GPS accuracy is roughly 20–25 m in the horizontal plane and 43 m in the vertical direction. The restricted PPS signal provides an accuracy of at least 22 m (typical values are 7–10 m) in the horizontal plane and 27.7 m in the vertical direction as well as coorinated universal time (UTC) time accuracy within 200 ns based on a reference signal from the US Naval Observatory.

GPS signals can be affected by multipath issues, where the radio signals reflect off surrounding terrain – buildings, canyon walls, hard ground, etc. This delay in reaching the receiver causes inaccuracy. A variety of receiver techniques, most notably narrow correlator spacing [20.20], have been developed to mitigate multipath errors. For long delay multipaths, the receiver itself can recognize the wayward signal and discard it. To address shorter delay multipaths due to the signal reflecting off the ground, specialized antennas may be used. This form of multipath is harder to filter out as it is only slightly delayed s compared to the direct signal, causing effects that are almost indistinguishable from routine fluctuations in the atmospheric delay.

20.5.3 Enhanced GPS

Wide-Area Augmentation System (WAAS)

The wide-area augmentation system (WAAS) is a supplementary signal that can be received by GPS receivers to improve their accuracy. WAAS increases the accuracy of horizontal position estimates from 10–12 m with GPS alone, to 1–2 m. The WAAS signal contains corrections for the GPS signal that reduce the effects of errors due to timing errors, satellite position corrections, and local perturbations due to variations in the ionosphere. These correction terms are estimated by ground-based stations at fixed and accurately known positions and uplinked to satellites which broadcast them to suitably enabled GPS receivers. The WAAS signal is only computed and available for North America, but similar correction signals are becoming available elsewhere as part of the standardization of satellite-based augmentation systems (SBAS). This includes Europe, where it is called the Euro Geostationary Navigation Overlay Service (EGNOS), and Japan and parts of Asia, where it is called the Multifunctional Satellite Augmentation System (MSAS). Further enhancements to GPS and WAAS, in the form of the Global Navigation Satellite System Landing System (GLS), are slated for completion in 2013.

Differential GPS

Differential GPS is a technique for correcting GPS signals by using a nearby GPS receiver located at a known accurately surveyed position. In fact, several variations on this basic idea exist and are also known under the general rubric of ground-based augmentation systems (GBAS). DGPS uses the same principles as WAAS but on a local scale without resorting to the use of satellite uplinks. The receiver at the known position computes the error in the GPS signal and transmits it to the nearby receiver at the unknown location. Since the error varies as a function of position on the earth, the effectiveness of the correction degrades with distance, typically with a maximum effective range of a couple of hundred miles. The method was especially desirable before the suspension of selective availability and the development of WAAS (which can be viewed as a form of DGPS). In the USA and Canada, a network of ground-based DGPS transmitters are in place, sending signals using radio frequencies between 285 kHz and 325 kHz. Commercial DGPS solutions akin to WAAS also exist.

Receiver Autonomous Integrity Monitor (RAIM)

Receiver autonomous integrity monitoring (RAIM) is a technique by which multiple pseudorange measurements (i. e., pose estimates) are obtained using different combinations of satellites. If inconsistent measurements are obtained, it indicates that a failure of some sort has taken place in the system. A position fix using at least five satellites is needed to detect such an error while at least six satellite are needed to exclude the data from a single bad satellite and still obtain a reliable estimate.

20.5.4 GPS Receivers and Communications

GPS receivers are classified according to their performance and cost. The best receivers are referred to as *geodetic grade* with economical models referred to as *resource grade* or *recreational*. In general the costs of these different models vary by several orders of magnitude, but the gap in performance is gradually narrowing.

Receivers come in two types: code phase and carrier phase. Code phase receivers use the satellite navigation message part of the data stream to provide the ephemeris data and produce real-time output. There is a delay for them to lock into the satellites, but then they produce output continuously without an initial position estimate. The C/A signal is a 1023 pseudorandom *noise* (PRN) bit string with a known key. The actual pseudorange data is determined by finding the phase offset of this bit string.

Carrier phase receivers, on the other hand, use the phase of the raw GPS signal rather than the embedded (digital) C/A signal. The L1 and L2 signals have wavelengths of 19 and 24 cm, respectively, and good-quality phase measurements allow horizontal positioning accuracies on the order of millimeters. These measurements, however, only provide relative position information within a neighborhood of some tens of kilometers.

Serial Protocols

Consumer GPS devices almost universally support some variant of the National Marine Electronics Association (NMEA) protocol, a serial protocol often transmitted using RS-232 wiring. Several variants of the protocol exist, but NMEA 0183 is the most commonly supported while NMEA 2000 supports higher data rates. While the protocol is proprietary and the official specification can only be purchased from the National Marine Electronics Association, there are several open-source descriptions of the protocol that have been reverse engineered.

The protocol supports an ASCII mode of communications based on a *talker* (the GPS receiver) and one or more *listeners* (computers) which receiver simple protocol strings called sentences. Anecdotal evidence suggests that there are ambiguities in the protocol that can lead to difficulties in assuring compatibility between devices (this suggestion is necessarily made without inspecting the proprietary documentation, which might preclude its description here).

20.6 GPS–IMU Integration

Although GPS offers the promise of high-resolution positioning information on or about the surface of the Earth, it does not solve all of the problems associated with robot pose estimation. First, it does not directly obtain information about vehicle orientation. To determine the orientation of the vehicle yaw, and for many vehicles pitch and roll, must be estimated by either differentiating the GPS signal or by integration with other sensors such as compasses, gyrocompasses, and IMUs. Second, GPS receivers are generally unable to provide continuous independent estimates of position. Estimates are only available at distinct time instances with (for inexpensive receivers at least) considerable delays between measurements. A continuous estimate of pose requires estimation of pose between GPS readings. Finally, it is not always possible to obtain a GPS fix. Local geography (e.g., mountains, buildings, trees) or an overhead cover that is opaque to radio signals (e.g., indoors, underwater) can block the signal entirely. Integration of a GPS receiver with another sensor technology (often an IMU) can be used to deal with these issues, at least for short periods of time.

The process of integrating GPS and IMU data is typically expressed as a Kalman filter estimation process (Sect. 25.2.3). Essentially the IMU data is used to bridge between solid GPS measurements and is combined in a least-squares optimal sense with the GPS data when both are available. Given the complementary nature and true independence of the two sensors, a wide range of commercial packages have been developed to integrate GPS and IMU data (see [20.21] for an example).

20.7 Further Reading

Odometry

Many general robotics books, including [20.2] and [20.22], provide considerable information on vehicle odometry and derive odometry equations for standard vehicle designs.

Gyroscopic Systems

Everett's book [20.23] provides a review of various sensor technologies including gyroscopic systems and accelerometers. Interesting historical documentation on the gyrocompass and its inventor can be found in *Hughes*'s book [20.5].

Accelerometers

Everett's book [20.23] provides a review of various sensor technologies, including gyroscopic systems and accelerometers.

GPS

Considerable details on the theory and implementation of GPS systems can be found in *Leick*'s book [20.24]. See also [20.25]. Details of various approaches to GPS-INS integration can be found in [20.26] and [20.27]

20.8 Currently Available Hardware

Although the specific models listed below are likely to have a short shelf life, the list of contacts may be a good starting point for the identification of specific inertial sensing devices.

Gyroscopic Systems

- KVN DSP-3000 Tactical Grade Fiber Optic Gyro (FOG).
  KVH Industries Inc., 50 Enterprise Center,
  Meddletown, RI 02842-5279, USA
- Fiber Optic Gyroscope HOFG-1(A).
  Corporate Headquarters Hitachi Cable Ltd.,
  4-14-1 Sotokanda, Chiyoda-ku, Tokyo 101-8971, Japan
- Rate Gyroscope CRS03,
  Silicon Sensing Systems Japan Ltd.,
  1-10 Fusocho (Sumitomo Precision Complex),
  Amagaxaki, Hyogo 660-0891, Japan

Accelerometers

- Accelerometer FA 101,
  A-KAST Measurements and Control Ltd., 1054-2 Centre St. Suite #299,
  Thornhill, ON, L4J 8E5, Canada
- ENDEVCO MODEL 22,
  Brüel and Kjær,
  DK-2850 Naerum, Denmark

IMU Packages

- µIMU,
  MEMSense, 2693D Commerce Rd.,
  Rapid City, SD 57702, USA
- IMU400 MEMS Inertial Measurement Unit,
  Crossbow Technology Inc.,
  4145 N. First St., San Jose, CA 95134, USA
- IntertiaCube3, (3DOF IMU),
  Intersense,

36 Crosby Dr, #15,
Bedford, MA 01730, USA

GPS Components

- Garmin GPS 18
  Garmin International Inc.,
  1200 East 151st St.,
  Olathe, KS 66062-3426, USA
- Magellan Meridian Color
  Thales Navigation
  471 El Camino Real,
  Santa Clara, CA 95050-4300, USA
- TomTom Bluetooth GPS Receiver
  Rembrandtplein 35,
  1017 CT Amsterdam,
  The Netherlands

References

20.1 Vitruvius: *Vitruvius: The Ten Books on Architecture* (Harvard Univ. Press, London 1914), English translation by M. Morgan

20.2 G. Dudek, M. Jenkin: *Computational Principles of Mobile Robotics* (Cambridge Univ. Press, Cambridge 2000)

20.3 E. Maloney: *Dutton's Navigation and Piloting*, 14th edn. (US Naval Inst. Press, Annapolis 1985)

20.4 J.G.F. Bohnenberger: Beschreibung einer Maschine zur Erläuterung der Gesetze der Umdrehung der Erde um ihre Achse, und der Veränderung der Lage der Letzteren, Tübinger Bl. Naturwiss. Arzneikunde **3**, 72–83 (1817), (in German)

20.5 T. P Hughes: *Elmer Sperry: Inventor and Engineer* (The Johns Hopkins Univ. Press, Baltimore 1993), Reprint Edition

20.6 H.W. Sorg: From Serson to Draper – two centuries of gyroscopic development, Navigation **23**, 313–324 (1976)

20.7 W. Davenport: *Gyro! The Life and Times of Lawrence Sperry* (Scribner, New York 1978)

20.8 J.F. Wagner: From Bohnenberger's machine to integrated navigation systems, 200 years of inertial navigation, Photogramm. Week **05**, 123–134 (2005)

20.9 R.P. Feynman, R.B. Leighton, M. Sands: *The Feynman Lectures on Physics* (Addison-Wesley, Reading 1963)

20.10 Navpers 16160: The Fleet Type Submarine. Produced for ComSubLant by Standards and Curriculum Division Training, Bureau of Naval Personnel, United States Navy (1946)

20.11 G. Sagnac: L'ether lumineux demontre par l'effect du vent relatif d'ether dans un interferometre en rotation uniforme, C. R. Acad. Sci. Paris **157**, 708–710 (1913), (in French)

20.12 G. Sagnac: Sur la preuve de la realitet de l'ether lumineaux par l'experience de l'interferographe tournant, C. R. Acad. Sci. Paris **157**, 1410–1413 (1913), (in French)

20.13 F. Harress: Die Geschwindigkeit des Lichtes in bewegten Körpern, Dissertation (Friedrich-Schiller-Universität, Jena 1912) (in German)

20.14 A.A. Michelson, H.G. Gale: The effect of the earth's rotation on the velocity of light, J. Astrophys. **61**, 140–145 (1925)

20.15 S. Ekeziel, H.J. Arditty: *Fiber-Optic Rotation Sensors*, Springer Ser. Opt. Sci., Vol. 32 (Springer, Berlin, Heidelberg 1982)

20.16 G.E. Stedman: Ring-laser tests of fundamental physics and geophysics, Rep. Prog. Phys. **60**, 615–688 (1997)

20.17 D. Mackenzie: From the luminiferous ether to the Boeing 757: A history of the laser gyroscope, Technol. Cult. **34**, 475–515 (1993)

20.18 E. Foxlin, M. Harrington, Y. Altshuler: Miniature 6-DOF inertial system for tracking HMDs, Proc. SPIE **3362**, 214–228 (1998)

20.19 M. Mostafa: History of inertial navigation systems in survey applications, J. Am. Soc. Photogramm. Remote Sens. **67**(12), 1225–1227 (2001)

20.20 A.J. Van Dierendonck, P. Fenton, T. Ford: Theory and performance of narrow correlator spacing in a GPS receiver, Navigation J. Inst. Navigation **39**, 265–283 (1992)

20.21 J. Rios, E. White: Low cost solid state GPS/INS package, Proc. Inst. Navigation Conf. (2000)

20.22 R. Siegwart, I.R. Nourbakhsh: *Introduction to Autonomous Mobile Robots* (MIT Press, Cambridge 2004)

20.23 H.R. Everett: *Sensors for Mobile Robots: Theory and Application* (Peters, Wellesley 1995)

20.24 A. Leick: *GPS Satellite Surveying* (Wiley, New York 2004)

20.25 P. Misra, P. Enge: *Global Positioning System: Signals, Measurements, and Performance*, 2nd edn. (Ganga-Jamuna, Lincoln 2006)

20.26 J. Farrell, M. Barth: *The Global Positioning System and Inertial Navigation* (McGraw-Hill, New York 1989)

20.27 M.S. Grewal, L.R. Weill, A.P. Andrews: *Global Positioning Systems, Inertial Navigation, and Integration* (Wiley-Interscience, New York 2007)