JASS 2006: Mechatronics - Foundations and Applications

Position Measurement in Inertial Systems

Christian Wimmer

April, 2006

Abstract

In this paper the principles of position and orientation measurement are explained and its implementation on an Initial Measurement Unit (IMU) are show.
The first part is about motivation for position measurement, dealing with industrial applications. Then we proceed to mathematical basics of attitude representation and figure out the advantages and weaknesses of the specific types of representation. Next is to give a brief of modern sensor technology. At last we show how to enhance the performance of a strapdown Inertial Measurement Unit using Kalman filtering.

1 Motivation

At present time the exact determination of position and orientation of technical systems plays an important role. Especially for robustly controlling the behavior of mechatronic systems a reliable position measurement system is required. Since the early years of the past century inertial navigation has made a great evolution. First there have only been military applications propulsing inertial navigation technology, but nowadays there is also a great variety of civil and industrial products using inertial navigation units (IMU). The development of inertial measurement systems in recent years has been characterized by the gradual move form stable gimbaled platform to stapdown technology, bringing a much more greater comfort to handle the systems, but also requiring better algorithms.
In the following, we want to give some representing examples:

- *the biped walking machine 'Johnnie':* the humanoid heavily depend on the performance of its position measurement sensors in order to keep balance and to get orientation. We brought on a strapdown IMU enabling him to exactly determine his position in 3D.
- *automotive industry:* IMUs are applied to measure the topology of test routes and to analyse drive dynamics.
- *aeronautic and space products:* helicopters, airplanes and space shuttles need highly accurate position measurement systems in order to keep control of its movement or to run an autopilot system. Of course also GPS navigation is used to determine the actual position, but since its resolution is not exact enough complementary measurement systems have to be installed.
- *military applications:* as already mentioned above, they have been the origin purpose of inertial navigation. First of all, strategic (for instance: ICBM, CM) or tactical missiles require position determination for navigation as well as torpedoes and jets. In recent times the development autonomous navigating systems, like drones is furthered.

- *maritime systems:* inertial navigation systems are not only used on ships for guidance, but even to stabilize and balance landing platforms for helicopters. In case of marine applications, such systems may be required to provide navigation data of high accuracy over periods of weeks or months.
- *industrial robotic systems:* many robots for maintenance (e.g. for pipelines) need guidance systems for position determination.

2 Basic principles of inertial navigation systems

2.1 Basic types of inertial navigation systems

Inertial measurement systems, also known as inertia measurement systems, detect the accelerations acting along their sensitive axes. Integrate the output once, you have velocity, integrate again, and you have position along the accelerometer's axis. As a body in 3D has 6 degrees of freedom, we need 3 sensitive axes standing perpendicular to each other and 2 sensors for each axis, one for acceleration (accelerometer) and one for rotation (gyroscope).

Basically, there exist two different types of inertial measurement technologies. Firstly, the gimbaled platform technology being the older one and still in use in many applications, which require high accuracy for a long period of time (i.e. Voyager space craft or submarines). The system basically consists of all 6 sensors being mounted on a platform and arranged perpendicular to each other as mentioned above. The platform is suspended in a set of 3 gimbals, that are gyro-stabilized to maintain the direction when the vehicle manoeuvres (Kardan rotation). The gyros are used as sensing elements in null-seeking servos, with the output of each gyro connected to a servo-motor driving the appropriate gimbal, thus keeping the gimbal in a constant orientation in inertial space. This leading to the fact, that the rotational motion of the IMU is completely decoupled from the vehicle's. But great efforts have to be done to realize the complex mechanisation of the Kardan platform.

But as microcomputers and gyroscopes with large dynamic ranges came up, the second and nowadays most important inertial measurement technology, the stapdown principle, could be realized. They are much more smaller and less complex than their predecessors, although gimbaled technology is still in use in some specific high performance applications. They consist of only one platform including all six sensors and is directly strapped down on a vehicle. Again the sensors are arranged perpendicular following an orthogonal axis system.

2.2 Reference frames

Fundamentally, a number of cartesian coordinate systems has to be defined precisely. Each system is a right handed coordinate frame. Due to the purpose a lot of different frames are conceivable, here are only some of the most common:

- *The inertial frame (i-frame):* has its origin at the center of the earth and non rotating axis according to the stars. All of its axis are time invariant, although the its center is translationally moving, as the earth turns.
- *the earth frame (e-frame):* has its origin at the earth's center and axis, that are fixed with respect to the earth.
- *the navigation frame (n-frame):* has its origin at the location of the navigation system and its axis aligned with the directions of north, east and the local vertical (down).
- *the wander azimuth frame (w-frame):* may be used to avoid the singularities in the computation which occur at the poles of the navigation frame. Like the n-frame it is locally level.
- *the body frame (b-frame):* is a local body-fixed axis system, which is aligned with the roll, pitch and yaw axes of the vehicle containing the navigation system.

According to the chosen frame, a suitable frame mechanisation has to be found. For the present, all measured values are received in the body coordinate system. The navigation equation may be solved in any one of a number of reference frames. But if the chosen frame rotates, the theorem of Coriolis has to be considered (Euler derivatives).

$$ {}_b\boldsymbol{v}_{P\,abs} = \boldsymbol{A}_{bi}\,{}_i\boldsymbol{v}_{P\,abs} = \frac{\mathrm{d}}{\mathrm{d}\,t}\left({}_b\boldsymbol{r}_{OP}\right) + {}_b\boldsymbol{\omega}_{ib} \times {}_b\boldsymbol{r}_{OP} \tag{1} $$

Where ${}_b\boldsymbol{\omega}_{ib}$ is the turn rate of the body frame with respect to the i-frame. The accelerometers usually provide a measurement of specific force in a body fixed axis set, denoted ${}_b\boldsymbol{f}$. In order to navigate it is necessary to resolve the components of the specific force in the chosen reference frame.

$$ {}_i\boldsymbol{f} = \boldsymbol{A}_{ib}\,{}_b\boldsymbol{f} \tag{2} $$

$$ \dot{\boldsymbol{A}}_{ib} = \boldsymbol{A}_{ib}\,{}_b\boldsymbol{\Omega}_{ib} \tag{3} $$

Where ${}_b\boldsymbol{\Omega}_{ib}$ is a skew symmetric matrix formed from the elements of the vector ${}_b\boldsymbol{\omega}_{ib}$ representing the turn rate of the body with respect to the i-frame as measured by the gyroscopes, and $\boldsymbol{A}_{ib}$ is the 3 x 3 direction cosine matrix. To give a short example of a frame mechanisation, the inertial frame is chosen for representation. The navigation equation reads as follows:

$$ {}_i\dot{\boldsymbol{v}}_P = \boldsymbol{A}_{ib}\,{}_b\boldsymbol{f} - {}_i\boldsymbol{\omega}_{ie} \times {}_i\boldsymbol{v}_P + {}_i\boldsymbol{g}_l \tag{4} $$

Where ${}_i\boldsymbol{g}_l$ is the result vector form gravitation and centripetal acceleration, since the two components can not be separately distinguished and ${}_i\boldsymbol{v}_P$ denotes the vehicle's speed with respect to the Earth, the ground speed, in inertial axes resolution.

2.3 Strapdown attitude representations

The attitude of the body frame, which is required to resolve the specific force measurements into the reference frame, may be defined in a number of different ways. Here, only the most common representations are listed:

- *Direction cosine matrix:* this is a 3 x 3 matrix denoted by $\boldsymbol{A}_{ib}$, the columns of which represent unit vectors in body axes projected along the reference axes. The element in the i-th row and the j-th column represents the cosine of the angle between the i-th axis of the reference frame and the j-th axis of the body frame. The rate of change of $\boldsymbol{A}_{ib}$ with time is given by Eq. 3. For computational calculation this kind of attitude representation is probably the most common, because of its simple handling and the fact that it doesn't contain any singularities.

- *Quaternions:* this type of attitude representation is based on the idea that a transformation from one coordinate frame to another may be done by a single rotation about a vector $\boldsymbol{\mu}$ defined with respect to the reference frame. The quaternion $\boldsymbol{q}$, is a four element vector, the elements of which are functions of this vector and the magnitude of the rotation.

$$p = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} \cos \mu/2 \\ (\mu_x/\mu) \sin \mu/2 \\ (\mu_y/\mu) \sin \mu/2 \\ (\mu_z/\mu) \sin \mu/2 \end{pmatrix} \quad (5)$$

$$p = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d \quad (6)$$

  Since quaternion representation doesn't have any singularities it is also very popular, despite they are not very graphic. Operations like addition or multiplication are similar to two parameter complex notation.

- *Euler angles:* a transformation from one coordinate frame to another is defined by three successive rotations about different axes taken in turn.The Euler angle representation is perhaps one of simplest techniques in term of physical appreciation. A new coordinate frame may be expressed as follows, the three rotation are performed in turns: rotate through angle ψ about reference z axis, rotate through angle θ about new y axis, rotate through angle ϕ about new x axis. As mentioned above, this corresponds to the angles measured by a angular pick off between a set of three gimbals in a stable platform inertial navigation system.

$$\boldsymbol{C}_{IK}=\begin{pmatrix}\cos\theta\cos\psi & -\cos\phi\sin\psi+\sin\phi\sin\theta\cos\psi & \sin\phi\sin\psi+\cos\phi\sin\theta\cos\psi\\ \cos\theta\sin\psi & \cos\phi\cos\psi+\sin\phi\sin\theta\sin\psi & -\sin\phi\cos\psi+\cos\phi\sin\theta\sin\psi\\ -\sin\theta & \sin\phi\cos\theta & \cos\phi\cos\theta\end{pmatrix} \quad (7)$$

It is useful to use direction cosine and quaternion representation for intern calculations and Euler angles for extern representation.

3 Sensor technology

3.1 Accelerometers

Inertial navigation relies on measurement of the acceleration which can be integrated successively to provide estimate of changes in velocity and position. Accelerometers measure both, the inertial force and the gravitation component, which have to be separated in order to obtain only acceleration. Since it is not practical to determine the acceleration of a vehicle by measuring the total force acting upon it, it is possible to measure the force acting on a small mass contained within the vehicle which is constrained to move with the vehicle. A simple sensor could be thought as a spring-mass structure packed in a small container.

3.2 Gyroscopes

To get the turn rates, two different types of sensors, each based on another physical effect can be used. Vibratory gyroscopes, basing in the principle of Corriolis, are the most common, although not very accurate. If a body moves along a direction while rotating about an axis perpendicular to that direction, an acceleration in a third direction orthogonal to the two others acts upon it. Usually, the velocity is generated by an harmonically vibrating piezo crystal. Although huge noise and temperature depending drifts occur, this is a very cheap technology. For low bandwidth purposes this aspect doesn't matter as well

as the fact, that the Unit is also negatively influenced by body translational acceleration, vibrations and the body's acoustic noise.
For more precise application optical sensors are required, basing on the so called Sagnac effect. Light beam coming from a super luminiszenz diode (SLD) moves through a beam splitter, dividing the beam. After having moved through a fiber optic cable coil, each in opposite direction, the effective path length difference between the two counterpropagating beams are detected. If the interferometer rotates the time of each light beam to pass around the circumference is modified.This is because of the motion of the beam splitter during the time taken for the light to pass around the ring. In case of using a laser instead of the SLD, which is so called ring laser gyroscope (RLG), because of the effective path length, the laser's frequency changes. Therefore, a direct digital angular signal is obtained.

4 Sensor fusion and Kalman filtering

4.1 Combination of independent estimates

The Kalman filtering process combines two or even more independent estimates of a variable in order to form a weighted mean value. For this reason, a higher order of accuracy could be achieved. Consider the single dimension case, in which two independent estimates x_1, provided by updating a previous best estimate in accordance with the known equations of motion and x_2, which is obtained by a measurement. With σ_1^2 being the covariance of estimate x_1 and σ_2^2 of x_2, for the weighted mean value follows:

$$x = x_1 - w(x_1 - x_2) \tag{8}$$

$$\overline{x} = E(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2, for\ n \leq \infty \tag{9}$$

$$\overline{x} = E(x) = w_1 E(x_1) + w_2 E(x_2) \tag{10}$$

Where $w = w_2$ ($w_1 = 1 - w$) is the weighting factor of quantity two. In multidimensional case in which x_1 and x_2 are n-dimensional vectors, the equation changes to:

$$\overline{x} = x_1 - W(x_1 - x_2) \tag{11}$$

Here W stands for a n x n weighting matrix. The best estimate of x is obtained when W is selected to minimise the covariance of $\overline{x}$. In practical, the dimensions of the two estimates are not equal. A set of m measurements may be provided, denoted y_2, where y_2 is only related to some of the elements of x:

$$y_2 = Hx_2 \qquad H \in \Re^{m \times n} \tag{12}$$

By definition, the covariance P of $\overline{x}$ for multidimensional case is given by:

$$P = E\left\{[\overline{x} - E(\overline{x})]\,[\overline{x} - E(\overline{x})]^T\right\} \tag{13}$$

With the above given measurement equation (Eq. 12) for y_2 the equation for $\overline{x}$ by replacing $W = KH$ reads as follows:

$$\overline{x} = x_1 - KH(x_1 - x_2) = (I_K H)x_1 + Ky_2 \tag{14}$$

Now, applying the formula for the covariance P for multidimensional case, we obtain:

$$\begin{aligned} P =& E\left\{[(I - KH)x_1 + Ky_2 - (I - KH)E(x_1) - KE(y_2)]^1\right. \\ & \left.[(I - KH)x_1 + Ky_2 - (I - KH)E(x_1) - KE(y_2)]^T\right\} \end{aligned} \tag{15}$$

And since x_1 and y_2 are uncorrelated, this reduces to:

$$\begin{aligned} P &= (I - KH)E\left\{[x_1 - E(x_1)]\,[x_1 - E(x_1)]^T\right\}(I - KH)^T \\ &+ KE\left\{[y_2 - E(y_2)]\,[y_2 - E(y_2)]^T\right\}K^T \\ &= (I - KH)P_1(I - KH)^T + KRK^T \end{aligned} \tag{16}$$

In which P_1 denotes the covariance of estimate x_1, such as R denotes the covariance of estimate x_2.

Now we are looking for K matrix, such that P minimises in the sense, that the diagonal elements of P, the variances of x, are minimised. As this is given in [1], K reads as:

$$K = P_1 H^T \left[H P_1 H^T + R \right]^{-1} \tag{17}$$

in which R denotes the variance of y_2.
Under such conditions, the best estimate of x is given by:

$$\overline{x} = x_1 - K \left[H x_1 - y_2 \right] \tag{18}$$

$$P = P_1 - K H P_1 \tag{19}$$

Note, that the quantity x_1 can not only be obtained by the motion equation, but also by a process called sensor fusion, which the both quantities obtained for instance by the long-term signal of the accelerometers and the short-term signals of the gyroscopes are merged together in. Even GPS, if provided, can be used to infuse further information into the system. This is probably the most common case.

4.2 Kalman filtering

A Kalman filter is simply an optimal recursive data processing algorithm. There are many ways of defining optimal, dependent upon the criteria chosen to evaluate performance. The Kalman filter is optimal with respect to virtually any criterion that makes sense. One aspect of this optimality that the Kalman filter incorporates all information that can be provided to it. It processes all available measurements, regardless of their precision , to estimate the current value of the variables of interest, with use of (1) knowledge of the system and measurement device dynamics, (2) the statistical description of the systems noises, measurements errors, and uncertainty in the dynamics models, and (3) any available information about the initial conditions of the variables of interest.
The dynamic behavior of a linearised system may be expressed by a first order differential equation (state representation):

$$\dot{x} = F x + G u + D w \tag{20}$$

Here, $x(t)$ denotes the system's state vector, $u(t)$ is the deterministic input vector and $w(t)$ is the system's noise. The output equation is expressed by:

$$y = Hx + v \tag{21}$$

Usually the noise vector $v(t)$ has zero-mean and is normally distributed, with spectral density R. The Kalman filter for this system described here seeks to provide the best estimates of the states $x(t)$. The measurement of the true system are compared with predictions of those measurements, derived from the latest best estimates of the states provided by the system model. The differences between the true and predicted measurements are fed back through a weighting matrix, the Kalman Gain matrix, to correct the estimated states of the model. The Kalman gain has to set in that way, to provide best estimates of the states in a least squares sense.
Since the measurements in practice are provided at discrete intervals of time the continuous differential equation given above has to be transformed to a discrete difference equation, as given in [1]:

$$x_{k+1} = \Phi_k x_k + \Gamma_k u_k + \Delta_k w_k \tag{22}$$
$$y_{k+1} = H_{k+1} x_{k+1} + v_{k+1} \tag{23}$$

The zero-mean, discrete noises, w_k for system's noise and v_{k+1} for measurement noise are characterized by covariance matrices Q_k and R_k respectively. In the process of filtering, two different sets of equations, the prediction process and the measurement update have to be distinguished. The best estimate of the state t_{k+1} at time t_k is:

$$x_{k+1} = \Phi_k x_k \tag{24}$$
$$P_{k+1} = \Phi_k P_k \Phi_k^T + \Lambda_k Q_k \Lambda_k^T \tag{25}$$

with P_{k+1} being the covariance at time t_{k+1}, predicted at time t_k.
When a new value y_{k+1} is obtained by measurement at time t_{k+1}, it is compared with the predicting value derived from the system's model. The measured value is then used to update the prediction value to

generate a best estimate. Hence the best estimate at time t_{k+1} is given by [3]:

$$x_{k+1/k+1} = x_{k+1/k} - K_{k+1}\left[H_{k+1}x_{k+1/k} - y_{k+1}\right] \tag{26}$$
$$P_{k+1/k+1} = P_{k+1/k} - K_{k+1}H_{k+1}P_{k+1/k} \tag{27}$$
$$K_{k+1} = P_{k+1/k}H_{k+1}^T\left[H_{k+1}P_{k+1/k}H_{k+1/k}^T + R_{k+1}\right] \tag{28}$$

Where H denotes the measurement matrix and K is the Kalman gain matrix. A Subscript followed by a slash denotes the time a value is calculated. If the system, which is focused nonlinear, the Kalman filter described above has to be updated to the extended Kalman filter, described below.

4.3 Extended Kalman filtering

The algorithm described in the previous section is only applicable to linear time invariant systems with Gaussian noise type disturbances. But as usually this condition is not appropriate, the algorithm has to be re-designed to the so called 'extended Kalman filter', in order to regain the optimal filtering. Since a filter of infinite dimension is not feasible in practice, we have to accept a suboptimal behavior and make its performance as close to optimal as possible. This concludes in predicting the system and its covariance matrix P over relatively short time intervals during which the conditions for linearity hold. This means that all matrices of the above linear system are no longer time invariant and have to be determinated for every time step. This is usually done by approximating the system about a nominal trajectory and linearisation by taylor series truncation. Let's consider a non linear continuous system given by:

$$\dot{x} = f(x, u) \tag{29}$$
$$y = h(x, u) \tag{30}$$
$$\tag{31}$$

also the trajectory is given by:

$$\dot{\tilde{x}} = f(\tilde{x}, \tilde{u}) \tag{32}$$

for small deviations, a truncated taylor series approximation can be assumed to be:

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \\ \frac{\partial f_n}{\partial x_1} & & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \quad (33)$$

$$x(t) = \tilde{x}(t) + \Delta x(t) \quad (34)$$

$$\dot{\tilde{x}} + \Delta\dot{x} = f(\tilde{x}, \tilde{u}) + \frac{\partial}{\partial x} f(\tilde{x}, \tilde{u})\Delta x + \frac{\partial}{\partial u} f(\tilde{x}, \tilde{u})\Delta u \quad (35)$$

The deviation term can be denoted in shorter form, using $A(t)$, $G(t)$, $H(t)$ and $D(t)$:

$$\Delta\dot{x} = A(t)\Delta x + G(t)\Delta u \quad (36)$$

$$\Delta y = H(t)\Delta x + D(t)\Delta u \quad (37)$$

As you can easily see, the system's matrices are time variant. In summary the following steps have to been done for each time interval to run the Kalman filter on the non-linear system:

1. Linearisation about the nominal trajectory, usually taken to be the latest estimate of the states.
2. Calculate the discrete system from the continuous.
3. Integrate the state prediction equations: the differential equations for the nominal trajectory and the deviations from it. Then adding the result.
4. Calculate the Kalman equations, to get best estimates of the deviations from the nominal trajectory. The correction summand, given by the Kalman gain multiplied by the measurement differences can added directly to the predicted state estimates.
5. Proceed to next time interval and return to 1.

This leads directly to the following discrete system representation of the non-linear system [3]:

$$x(t_{i+1}) = \Phi(t_{i+1}, t_i)x(t_i) + G(t_i)u_(t_i) + w(t_i) \quad (38)$$

$$y(t_i) = H(t_i)x(t_i) + v(t_i) \quad (39)$$

This equation expresses the error deviation from the system state $x(t_{i+1})$, in which the matrices Φ, G and H represent the vector derivatives. By using them instead of the timeinvatiant matrices of the linear filter, the Kalman Gain matrix for the non linear system is computed as given above.ead of the timeinvatiant matrices of the linear filter, the Kalman Gain matrix for the non linear system is computed as given above.

5 Conclusion

The basic principles of position measurement and navigation in inertial systems have been presented, in order to give a brief to this theme. The strapdown technology, which plays the most important role for the future, has been identified as the most sensitive because of its heavy addiction of precise attitude measurement. Therefore, high performance gyroscopes must be combined with accurate filter algorithms like the Kalman filter, in order to get further information from extern, improving the system's accuracy.

References

[1] D.H. Titterton, J.L. Weston, 'strapdown inertial navigation technology', *IEE radar, sonar, navigation and avionic series 5, 1997*

[2] E. v. Hinüber, 'Inertiale Meßsysteme mit faseroptischen Kreiseln', *iMAR Gmbh St. Ingbert, 2002*

[3] Z. Chen, 'Entwicklung und Implementierung eines Extended Kalman Filters zur Kompensation der Fehlerdynamik eines bioanalogen Inertialen Messsytems', *Semesterarbeit, TU München, Lehrstuhl für angewandte Mechanik, 2005*

[4] P. S. Maybeck, 'Stochastic Models, Estimation, and Control Volume 1', *Academic Press, New York San Francisco London, 1979*