Visual-Inertial Navigation Systems and Technologies

Jorge Alejandro Valdez-Rodríguez, Julio César Rodríguez-Quiñonez, Wendy Flores-Fuentes, Luis Roberto Ramírez-Hernández, Gabriel Trujillo-Hernández, Oscar Real-Moreno, Moisés J. Castro-Toscano, Jesús Elías Miranda-Vega, and Paolo Mercorelli

Abbreviations

CCW	Counterclockwise
CW	Clockwise
DCM	Direction cosine matrix
FOV	Field of view
GNSS	Global navigation satellite system
IMU	Inertial measurement unit
INS	Inertial navigation system
LSS	Laser scanner system
NED	North, east, down
NEU	North, east, up
OSV	Omnidirectional stereo video
OVINS	Omnidirectional visual-inertial navigation system
RMIS	Robot-assisted minimally invasive surgery

J. A. Valdez-Rodríguez · J. C. Rodríguez-Quiñonez (✉) · W. Flores-Fuentes
L. R. Ramírez-Hernández · G. Trujillo-Hernández · O. Real-Moreno · M. J. Castro-Toscano
Facultad de Ingeniería Mexicali, Universidad Autónoma de Baja California, Mexicali, Baja California, México
e-mail: valdez.jorge71@uabc.edu.mx; julio.rodriguez81@uabc.edu.mx; flores.wendy@uabc.edu.mx; luis.ramirez16@uabc.edu.mx; gabriel.trujillo@uabc.edu.mx; oreal@uabc.edu.mx; moises.castro@uabc.edu.mx

J. E. Miranda-Vega
Tecnológico Nacional de México/IT de Mexicali, Mexicali, Baja California, México
e-mail: elias.miranda@itmexicali.edu.mx

P. Mercorelli
Leuphana University of Lueneburg, Lueneburg, Germany
e-mail: paolo.mercorelli@leuphana.de

O. Sergiyenko (ed.), *Optoelectronic Devices in Robotic Systems*,
https://doi.org/10.1007/978-3-031-09791-1_6

SNR	Surgical navigation robot
SVS	Stereoscopic vision system
VINS	Visual-inertial navigation system
VIO	Visual-inertial odometry

1 Introduction

Navigation is an ancient activity of the human history, originated from the necessity of travel from a region to another in the search for a better environment, better resources, or better opportunities [57]. It is a task that not everyone is able to achieve; even the most experienced person can get lost in a trip and never arrive at the desired location. Therefore, a diversity of tools and methods are developed as an aid for travelers; those elements are known in navigation as references.

The use of references as support in navigation could be interpreted as an external object used for orientation. Landscapes are some of the oldest and effective references, but as the task becomes more complicated, the instrumentation used in navigation grows in precision and complexity.

However, navigation is not exclusive to travel; in the present days, a robot performing the specific task of placing objects in a diversity of defined points is part of the navigation problem. The possibilities of following a track, evading objects, and mapping a room are elements of navigation; they require a robot which has the capability of orientating in unknown environments.

For example, a pipeline inspection gauge is a system which has the capacity of performing navigation to inspect a gas or oil pipeline to locate and detect critical deformations [6]. The system propelled possesses a navigation system conformed of inertial sensors and/or GPS signals, allowing to track and enlist the places where a failure exists. Other tasks demand low error orientation as a result of the consequences in a poor navigation. For instance, the field of medicine requires manipulators capable of performing meticulous operation procedures in humans [61]. Teleoperated robot-assisted minimally invasive surgery (RMIS) is more common and an important part of medical surgery. RMIS systems require precision in their movements, sometimes to compensate the inexperienced movements of novice surgeons or simply to balance between high clinical importance and technical complexity.

Different resources are implemented as reference for INS, although a popular and continuously growing solution is the addition of visual references as cameras and laser scanners [46, 71]. Vehicles, such as cars, planes, and teleoperated RMIS, are technologies taking advantage of INS with visual sensors. The chapter presented discusses how INS are aided by visual references and the benefits gained in modern technologies.

2 VINS

Navigation is the science of maneuvering from one point to another using references to know the current position [8]. Tools, as maps or compasses, are commonly used when someone has in mind the word navigation, but more sophisticated devices are applied nowadays and the integration of two or more forms a "navigation system." Popular gadgets as cameras, instrumentation as accelerometers and gyroscopes, or systems as laser scanners are elements included in navigation systems.

The combination of two or more navigation references with navigation techniques, physics, and mathematical analysis receives the name of "navigation system" [42]. Navigation references are classified according to the coordinate reference frame, where devices with a fixed origin are known as "absolute references" [81]; otherwise, instruments with a relative coordinate reference frame where the position and attitude require to be constantly calculated over time are called "inertial references" [20].

A widely implemented instrument in inertial navigation is the inertial measurement unit (IMU) which is conformed by inertial instruments as accelerometers and gyroscopes. The combination of IMUs and mathematical navigation calculus is known as inertial navigation system (INS) [11] Fig. 1.

As a result of the mathematical integration required to obtain position and attitude, inertial sensors have an inherent bias that reduces precision in the navigation calculation. A common effect in INS is the presence of "drift," a deviation in the estimated position and orientation. Drift is an accumulative multifactorial error which predicts a different location from the actual body position as time elapses.

In order to reduce the drift error, a variety of solutions are recommended in literature to assist INS. Kalman filters help to improve the computational efficiency and diminish the error in navigation applying two weighting factors to compute a new estimation: the previous estimation in accordance with the known equations of motion and the obtained measurement from the IMU [16, 76]. The new estimation calculated by the Kalman filter increments the accuracy but is not capable of reducing all the drift error of an inertial sensor and computational miscalculations. It is possible to enhance the data accuracy of Kalman filter by complementing it with a different type of filters to correct computation.

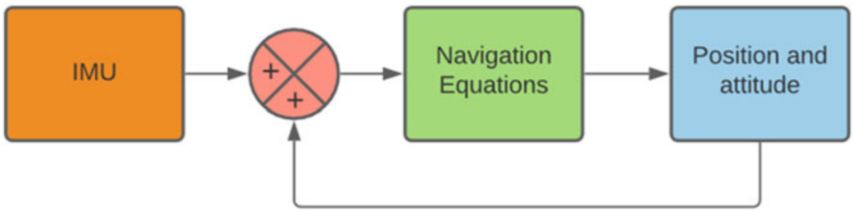


Fig. 1 Inertial navigation system

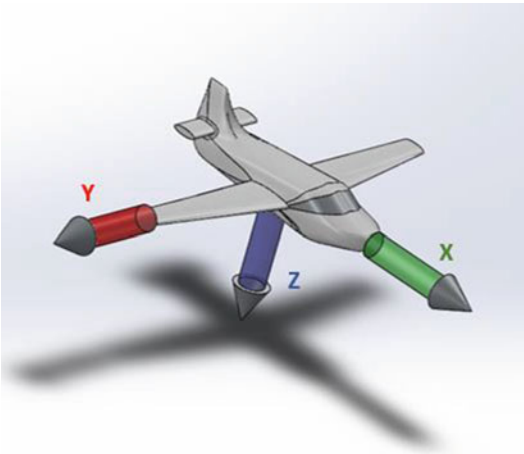
Another approach to reduce drift error in INS is the implementation of external navigation references such as GNSS (global navigation satellite systems), magnetometers (electronic compass), and visual systems as cameras to the INS. Visual sensors provide information about the environment where the body is located; however, they are sensitive to illumination conditions and motions [60]. An INS complemented with a visual system is called visual-inertial navigation system (VINS); the structure is an INS combined with one or more visual references as cameras, visual odometry [25], or laser scanner techniques.

The used approach in image-based navigation to complement INS is supplying and processing images using visual systems to provide accurate data of the surrounding environment or navigation object [59]. In addition, image-based navigation emulates the sense of orientation and navigation in human beings, allowing to determine object attitude and position in addition to object recognition [24].

In navigation is required a coordinate reference frame to express the position of a point in relation to some reference [43]. A coordinate reference frame is a Cartesian, right-handed axis set defined by a reference. Objects, the point of view, the Earth, and sensors are examples of references adopted to create a reference frame and with the aid of mathematical transformations are possible to translate from one coordinate reference frame to another.

One example of coordinate reference frame is the body frame. The body frame is the coordinate reference frame related to the vehicle or navigation object. The axes are related to the direction of the movements of the body, where **X** is the forward direction, **Y** is the right direction, and **Z** is related to the gravity pointing in the down direction (Fig. 2). In navigation, the body frame is assumed to have the axes in the same direction as the inertial frame to align the inertial sensors with the navigation vehicle. There is a wide variety of coordinate reference frames in navigation and

Fig. 2 Body frame



some of the related visual systems are going to be discussed further in the present chapter.

The translation from a coordinate reference frame to another is commonly solved through two different rotations in a plane, using methodologies such as the DCM and quaternion.

DCM is applied in a three-dimensional space R^3 where rotations are given through a coordinate angle called Euler angle [26, 62]. The rotations follow the right-handed coordinate frame rule where every rotation is defined, e.g., a CW rotation as positive and CCW rotation as negative.

A rotation in the Z axis is called yaw and the Euler angle is represented with ψ letter. The C_ψ rotation matrix is represented as follows:

$$C_\psi = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

The X axis rotation is represented by φ the Euler angle and is defined as roll. Eq. (2) shows the corresponding rotation matrix C_φ:

$$C_\varphi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix} \tag{2}$$

And for a rotation in the Y axis, the θ Euler angle is used and the rotation is named pitch. The C_θ rotation matrix is described below:

$$C_\theta = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \tag{3}$$

Therefore, for a succession of three rotations in each of the mentioned axis is created a DCM which represents a general translation from an A frame to a B frame:

$$C_{\mathrm{A}}^{\mathrm{B}} = \begin{bmatrix} \cos\theta\cos\psi & -\cos\varphi\sin\psi + \sin\varphi\sin\theta\cos\psi & \sin\varphi\sin\psi + \cos\varphi\sin\theta\cos\psi \\ \cos\theta\sin\psi & \cos\varphi\cos\psi + \sin\varphi\sin\theta\sin\psi & -\sin\varphi\cos\psi + \cos\varphi\sin\theta\sin\psi \\ -\sin\theta & \sin\varphi\cos\theta & \cos\varphi\cos\theta \end{bmatrix} \tag{4}$$

It is important to take into consideration that any translation between frames could perform two or more successive rotations involving Eqs. (1, 2, and 3). In this chapter are presented some examples applied for specific cases.

However, there exists another form to interpret navigation translations; this is by using quaternions [30, 41]. As the name suggests, a quaternion is formed by four elements, an "s" scalar value and a vector $\overrightarrow{v}$ conformed of three scalars representing an axis x, an axis y, and an axis z as shown below:

$$q = \begin{bmatrix} s \\ \overrightarrow{v} \end{bmatrix} = \begin{bmatrix} s \\ v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} q_s \\ q_x \\ q_y \\ q_z \end{bmatrix} \tag{5}$$

Therefore, every rotation in a quaternion is governed by the following equation:

$$q_{\mathrm{B} \leftrightarrow A} = \begin{bmatrix} q_s \\ q_x \\ q_y \\ q_z \end{bmatrix} = \begin{bmatrix} \cos \theta / 2 \\ \| \overrightarrow{e} \| \bullet \sin \theta / 2 \end{bmatrix} \tag{6}$$

where:

- $q_{B \leftrightarrow A}$ defines a translation from an A reference frame to a B reference frame and backward. Furthermore, the expression could represent a one-sided rotation if it is necessary.
- $\| \overrightarrow{e} \|$ vector represents a rotation in the axis of interest, which could be described as one of the following vectors shown in Eq. (7):

$$\| \overrightarrow{x} \| = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \| \overrightarrow{y} \| = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \| \overrightarrow{z} \| = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{7}$$

Additionally, it is possible for mentioned vectors $\| \overrightarrow{e} \|$ to represent a combination of two or three simultaneous rotations.

Hence, to rotate θ degrees a three-dimensional vector $\overrightarrow{v_a}$ and finalize in a $\overrightarrow{v_b}$ position, the equation is described as quaternion multiplication $\bigotimes$ as shown in Eq. (8):

$$\overrightarrow{v_b} = q_{B \leftarrow A} \bigotimes \begin{bmatrix} s \\ \overrightarrow{v_a} \end{bmatrix} \bigotimes q_{B \leftarrow A}{}^{-1} \tag{8}$$

Therefore, for every reference frame, different sequences of rotations are realized according to the planes and elements involved. Consequently, it is possible to convert the information from a DCM structure to a quaternion representation.

A quaternion is described by the diagonal elements of a DCM, where the equations applied are:

$$q_s = \sqrt{\frac{1}{4} \bullet (1 + C_{11} + C_{22} + C_{33})} \tag{9}$$

$$q_x = \sqrt{\frac{1}{4} \bullet (1 + C_{11} - C_{22} - C_{33})} \tag{10}$$

$$q_y = \sqrt{\frac{1}{4} \bullet (1 - C_{11} + C_{22} - C_{33})} \tag{11}$$

$$q_z = \sqrt{\frac{1}{4} \bullet (1 - C_{11} - C_{22} + C_{33})} \tag{12}$$

Thus, some elements of the DCM matrix required for one of the quaternion values may be equal to zero. Consequently, it is necessary to involve the elements in the other subdiagonals of the matrix [5].

The process followed indicates the evaluation of the recently computed q_s, q_x, q_y, and q_z, where the one with the greatest absolute value is selected. The other elements are recalculated accordingly to the selected value.

For a selected q_s, the other elements are estimated as:

$$q_x = \frac{C_{32} - C_{23}}{4 \bullet q_s} \tag{13}$$

$$q_y = \frac{C_{13} - C_{31}}{4 \bullet q_s} \tag{14}$$

$$q_z = \frac{C_{21} - C_{12}}{4 \bullet q_s} \tag{15}$$

If the determined greatest absolute value is q_x, then the equations for the other elements are:

$$q_s = \frac{C_{32} - C_{23}}{4 \bullet q_x} \tag{16}$$

$$q_y = \frac{C_{21} + C_{12}}{4 \bullet q_x} \tag{17}$$

$$q_z = \frac{C_{13} + C_{31}}{4 \bullet q_x} \tag{18}$$

When the greatest absolute value is q_y, the equations for the other values are:

$$q_s = \frac{C_{13} - C_{31}}{4 \bullet q_y} \tag{19}$$

$$q_x = \frac{C_{21} + C_{12}}{4 \bullet q_y} \tag{20}$$

$$q_z = \frac{C_{32} + C_{23}}{4 \bullet q_y} \tag{21}$$

And finally, if the greatest absolute value is q_z, the equations are determined as follows:

$$q_s = \frac{C_{21} - C_{12}}{4 \bullet q_z} \tag{22}$$

$$q_x = \frac{C_{13} + C_{31}}{4 \bullet q_z} \tag{23}$$

$$q_y = \frac{C_{32} + C_{23}}{4 \bullet q_z} \tag{24}$$

Furthermore, the elements in a quaternion can describe a DCM matrix if necessary [22]. The equation is described as:

$$C_A^B = \begin{bmatrix} q_s^2 + q_x^2 - q_y^2 - q_z^2 & 2 \bullet (q_x \bullet q_y - q_y \bullet q_s) & 2 \bullet (q_x \bullet q_z + q_y \bullet q_s) \\ 2 \bullet (q_x \bullet q_y + q_z \bullet q_s) & q_s^2 - q_x^2 + q_y^2 - q_z^2 & 2 \bullet (q_y \bullet q_z - q_x \bullet q_s) \\ 2 \bullet (q_x \bullet q_z - q_y \bullet q_s) & 2 \bullet (q_y \bullet q_z + q_x \bullet q_s) & q_s^2 - q_x^2 - q_y^2 + q_z^2 \end{bmatrix} \tag{25}$$

It is important to indicate that the information presented in the chapter to translate from one reference frame to another is expressed in DCM, but as previously described, it is possible to handle the rotations with quaternions.

Stereoscopic vision systems and laser scanner systems are two examples of visual systems applied in VINS. They are widely used in navigation applications along with popular technologies as LIDAR and are part of current methodologies. The aim of both systems is to provide absolute references to diminish the inherent drift error of INS using visual sensors as cameras or photoelectric sensors.

3 Stereoscopic Vision Systems

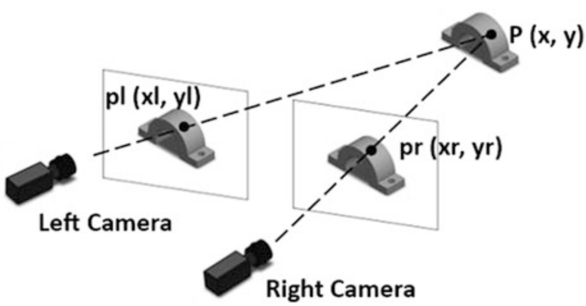
The cameras in navigation attempt to recreate the eyes' function, giving information about the surrounding environment. One single camera will only offer some information of the environment, but for terms of depth, surface shape, and curvature, it is necessary to add two or more cameras.

Stereoscopic vision systems, or SVS, acquire visual information from two or more cameras to obtain features of a specific scene [50]. SVS are portable systems with a wide field of view (FOV), capable of obtaining distance and object information, and also, SVS have advantages over other navigation devices as sonar and radar, because they do not require mechanical components and attain the pixels of the image at the same point in time [39, 80] (Fig. 3).

On the other hand, if SVS loss information in the image digitalization, there is a distortion in the lens or the system cannot find the corresponding points in the two (or multiple) images; they will not be available to achieve the triangulation process [51].

The integration of SVS to a VINS states a system capable of performing visual odometry, where two or more cameras work in conjunction with inertial references as an IMU to navigate in real time [21]. The SVS acts as an absolute reference for the INS, helping the inertial references with information about the surroundings to diminish drift and reduce the error in position and attitude.

Fig. 3 Stereoscopic vision system



For a navigation environment, SVS systems are subject to constant motion, fast dynamics, limited computation on board (for small vehicles or bodies), and the constant necessity of odometry [1]. In VINS, a SVS system must operate in a self-positioning configuration, where a camera is part of the body hardware and provides images of the current environment where the body is navigating [72].

In order to navigate implementing a VINS arrangement with an SVS as the visual sensor, the system acquires a set of images according to the number of cameras available. Therefore, it initiates the detection of significant points and recognizes geometries in the set of images, gets information of the surrounding, and then finds the similarities. Afterward, a pattern match process begins where points are localized in each image to subsequently identify the same points but in the previous iteration of the set of images. Finally, the system identifies the variations of the pixels in the sequence of the images to perform the estimation of the motion [48, 56].

Figure 4 shows a block diagram where an SVS process is integrated in an INS, resulting in VINS. The SVS process is followed by a block transforming the information to body reference frame, allowing the data to properly be compared with the INS dead reckoning and feedback the system.

Figure 4 shows that it is necessary to transform the data coming from the SVS process to a body reference frame. As other references, the cameras in a SVS own a reference frame according to their properties. A camera coordinate reference frame is shown in Fig. 5; the axis is aligned following the right-hand rule. In the presented frame, Z axis is pointing to the object or the environment in direction of the depth, which is the view of the camera; X and Y axes, on the other hand, follow the image axis.

As noted, the camera frame is formed by a different coordinate reference frame in comparison with the body frame (Fig. 2), a common situation when diverse references are integrated in any navigation system. Therefore, the axis is aligned through a mathematical transformation matrix to avoid misinterpretations during the navigation. A particular solution for multiple coordinate reference frames is rotating through the coordinate frames until arriving at the navigation frame, where the navigation is interpreted and related to a remarkable amount of coordinate reference

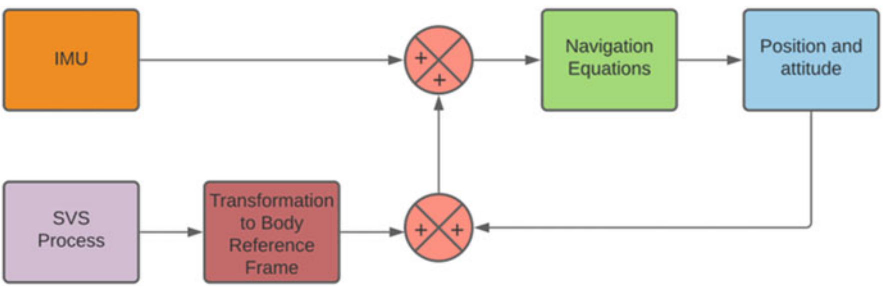


Fig. 4 SVS in an INS

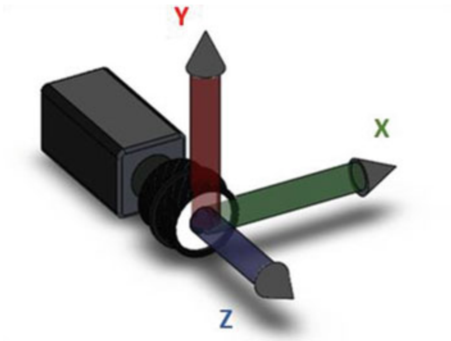


Fig. 5 Camera coordinate reference frame

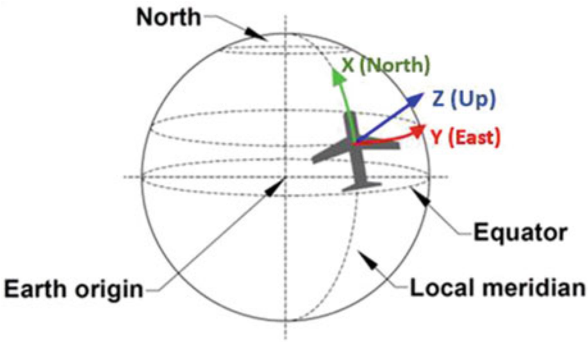


Fig. 6 Navigation frame

frames. The navigation frame is configurated with earth gravity and the cardinal points north and east. X axis is pointing to north; Y axis is aligned to east; and Z axis is pointing up or down, creating a NEU and a NED configuration. The arrangement is specified by the user (Fig. 6).

According to Wang et al. [73] and Veth [69], the camera is modeled after the camera perspective as showed in Fig. 7, where the presented frame shows the virtual frame in substitution of the focal plane to correct the inversion in the directions of the X and Y axes [74]. Also, Eq. (26) shows the line of sight vector from the camera pinhole in navigation coordinates (s^{n}):

$$s^{\mathrm{n}} = [x_i - x_c \quad y_i - y_c \quad - f] \tag{26}$$

The line of sight vector s^{n} is the difference between the image target location (i) and the camera position (c) in X, Y coordinates; x_i and y_i are the image coordinates and x_c and y_c are the camera coordinates. The vector also includes the camera's focal length f which is the distance between the camera and the image.

s_c^n, describes the cameras position transformed from camera frame to navigation frame. Therefore, the equation describes two DCM and considers the distance between the image target location and the camera position (27):

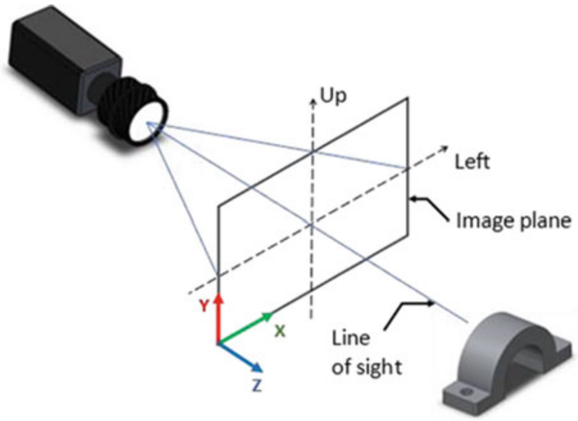


Fig. 7 Camera pinhole

$$s_{\mathrm{c}}^{\mathrm{n}} = C_{\mathrm{c}}^{\mathrm{b}} C_{\mathrm{b}}^{\mathrm{n}} s^{\mathrm{n}} \tag{27}$$

C_b^n is the DCM to transform from body frame to navigation frame. It is defined in Eqs. (28) and (29), where φ expresses the roll angle in X axis, θ is the pitch angle in Y axis, and ψ is the yaw angle in Z axis:

$$C_{\mathrm{b}}^{\mathrm{n}} = R_x(\varphi)\, R_y(\theta)\, R_z(\psi) \tag{28}$$

$$C_{\mathrm{b}}^{\mathrm{n}} = \begin{bmatrix} \cos\psi\cos\theta & \sin\psi\cos\theta & -\sin\theta \\ -\sin\psi\cos\phi + \cos\psi\sin\theta\sin\phi & \cos\psi\cos\phi + \sin\psi\sin\theta\sin\phi & \cos\theta\sin\phi \\ \sin\psi\sin\phi + \cos\psi\sin\theta\cos\phi & -\cos\psi\sin\phi + \sin\psi\sin\theta\cos\phi & \cos\theta\cos\phi \end{bmatrix} \tag{29}$$

$C_{\mathrm{c}}^{\mathrm{b}}$ DCM represent the camera to body transformation and involve rotations in Z and Y axes, transforming the attitude from the camera frame to the body frame (Eq. 30) [7]. The equation shows an azimuth angle expressed as α for Z axis and an elevation angle expressed as β for the rotation in Y axis:

$$C_{\mathrm{c}}^{\mathrm{b}} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\beta) & 0 & -\sin(\beta) \\ 0 & 1 & 0 \\ \sin(\beta) & 0 & \cos(\beta) \end{bmatrix}$$

$$C_{\mathrm{c}}^{\mathrm{b}} = \begin{bmatrix} \cos(\beta)\cos(\alpha) & \sin(\alpha) & -\sin(\beta)\cos(\alpha) \\ -\sin(\alpha)\cos(\beta) & \cos(\alpha) & \sin(\alpha)\sin(\beta) \\ \sin(\beta) & 0 & \cos(\beta) \end{bmatrix} \tag{30}$$

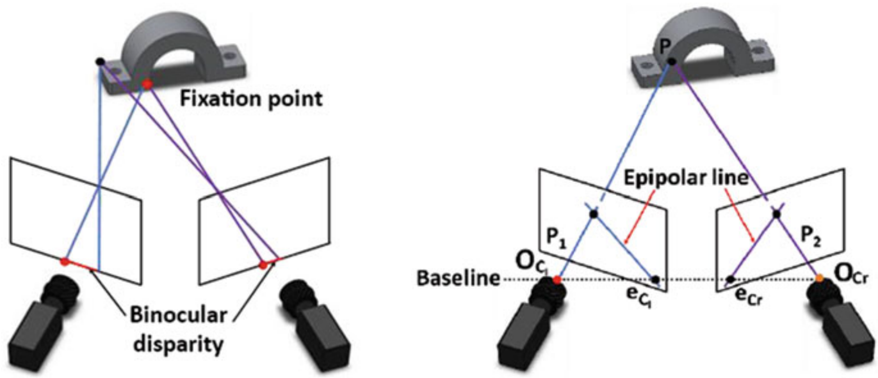


Fig. 8 Binocular disparity (**a**) and epipolar geometry (**b**)

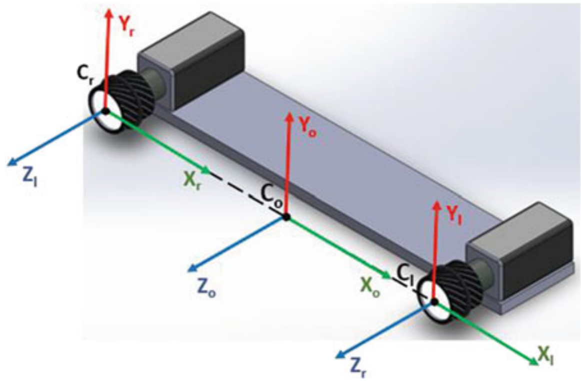


Fig. 9 Active stereo coordinate frame

As mentioned before, an SVS implements two or more cameras in order to estimate depth of objects and planes in a scene. The calculation of depth is produced by a binocular disparity created between the cameras [12]. Each camera possesses its own projection of the image, a triangulation between the points of each camera, and a point in the object of view creates an epipolar plane and determines a pair of epipolar lines in the two images, where the epipole point is the center of projection of the other camera [15, 29] (Fig. 8).

Consequently, to integrate a SVS in an INS, both cameras' frames are located in a midpoint frame where the two encounters. The active stereo coordinate reference frame C_{AS} describes the relation of left and right cameras, C_{l} and C_{r}, respectively; to an origin position, the calculation is made through a cross product as shown in Eq. (31) and Fig. 9 [31, 40].

$$C_{\mathrm{AS}} = (C_{\mathrm{r}} - C_{\mathrm{l}})\,(C_{\mathrm{o}} - C_{\mathrm{l}}) \tag{31}$$

The active stereo coordinate reference frame is located in the center and provides the data used in dead reckoning, where the C_{ASk} is compared with the $C_{\mathrm{ASk-1}}$ to compute a new attitude and position estimation.

4 Mobile Binocular Visual Inertial Odometry

Visual-inertial odometry (VIO) is a method employed in navigation to estimate motion using images acquired by camera sensors [64].

VIO systems are capable of functioning with a monocular camera; however, it is recommended to use a binocular camera to get better results in terms of environment recognition and motion due to its ability to perform depth determination [33]. Therefore, it is possible to implement a VIO system in vehicles or robotic manipulators to perform navigation tasks and work in conjunction with inertial instruments.

Vehicles as planes, drones, all-terrain mobile robots, or humanoid robots rely on binocular vision for navigation. For example, explorer vehicles take advantages of the cameras to travel across unknown environments and avoid collisions, or being positioned in a dangerous place [78].

All-terrain mobile robots are common in emergency situations in which there is no opportunity for persons to enter buildings. Binocular VIO aids not only to navigate through the place but also to provide opportunity to recognize and locate objects. A similar situation applies for drones where it is necessary to fly and identify landmarks [28] (Fig. 10).

However, VIO binocular mobile robots require obtaining non-blurry images to properly execute the process between frames and calculate the current distance and attitude [52]. Hence, VIO binocular mobile systems could be affected by the same surrounding environment they are exploring. Weather conditions, non-plain floors,

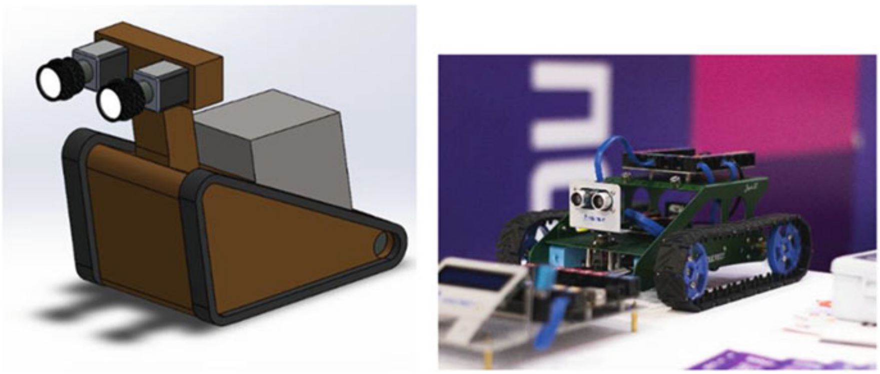


Fig. 10 SVS binocular rescue mobile vehicle (left) and binocular educational mobile vehicle (right) [10]

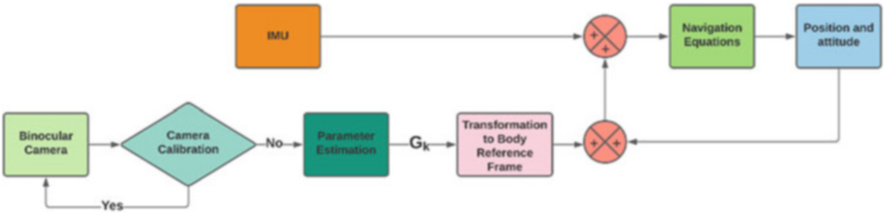


Fig. 11 Calibration and correction for blurry images

absence of light, or the cameras being directly affected by light are some of the drawbacks impeding the cameras to accurately get images.

VIO binocular mobile robots follow the majority of the steps in the process of VINS. However, in some systems, at the beginning a camera calibration phase and in a subsequent state a learning technique for parameter estimation through random samples are added [23, 66]. The methodologies mentioned work as a help to diminish the errors for the image taken by the binocular cameras.

Besides the feature point extraction of the image, the parameter estimation compares the current image with the previous one in order to obtain the current position and attitude of the navigation body or the relative motion T_k of the camera [4, 77]. Relative motion T_k is the computation of the current position and attitude of the navigation body, where every new data k obtained from the SVS process is concatenated with the previous data $k-1$. T_k is expressed as:

$$T_{k,k-1} = \begin{bmatrix} R_{k,k-1} & t_{k,k-1} \\ 0 & 1 \end{bmatrix} \tag{32}$$

where $R_{k,k-1}$ is a rotation matrix. T_k is a translation vector between frames taken at timesteps. Therefore, T_k is employed to determine a global estimate G_k to transform the information to the body reference frame in the particular case of VINS, as shown in Fig. 11 and afterward, compare the information with inertial instruments as IMU. G_k is obtained with the previous G_{k-1} and the relative motion T_k referenced at the initial frame G_0 at $k = 0$:

$$G_k = G_{k-1} T_k \tag{33}$$

5 Omnidirectional Visual-Inertial Navigation Systems

OVINS are navigation systems build with two omnidirectional cameras or two rotating cameras; their purpose is to measure depth in the horizontal plane of the cameras additionally to horizontal and vertical distances. In navigation, the device

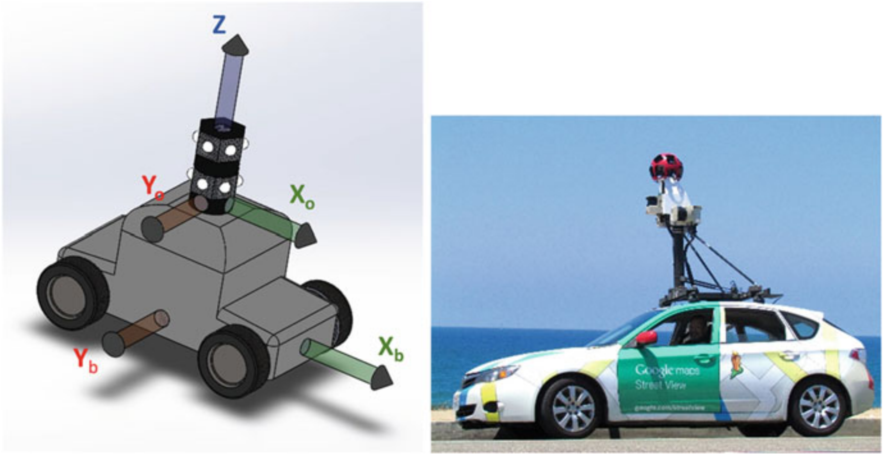


Fig. 12 OVINS: Two omnidirectional cameras in a vehicle to calculate depth (left) and Google Maps car (right) [9]

Fig. 13 OVINS detecting points in two omni-images

employed must be able to record omnidirectional stereo video (OSV) and create panoramic images to estimate distance and attitude in real time (Fig. 12)

In a city, OVINS allow INS to be aware of the different obstacles a mobile vehicle may encounter. From small places as malls or amusement parks to a bigger environment as the street, OVINS have the possibility to determine the distance between the mobile vehicle and the object (Fig. 13).

In OVINS, the FOV they possess is bigger than the one proportioned in conventional cameras; it allows the possibility to perform navigation tasks in a large environment. UAV as drones perform specific tasks such as recognition and reconstruction of the surroundings where the mobile vehicle is navigating. The

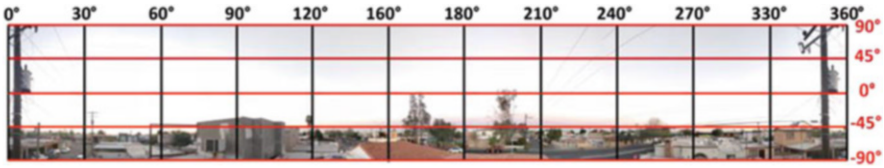


Fig. 14 Panoramic image and axis

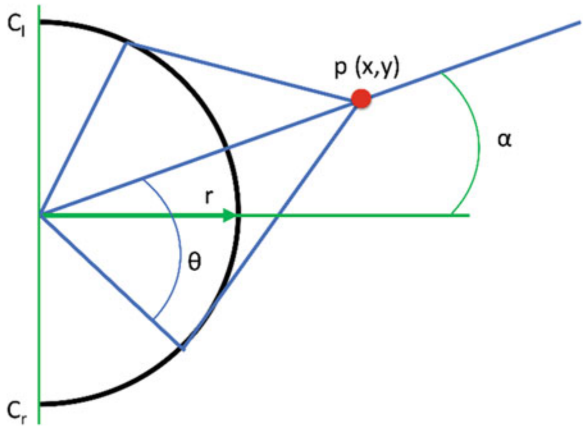


Fig. 15 Left and right camera FOV over a point $p(x, y)$

OVINS also realize odometry work with the images, defining coordinates for the points demonstrated in the captured spherical images [49].

OVINS takes panoramic images of the surroundings, where the image shows the horizontal and vertical axis through the pixels. A panoramic image is able to display between 0° and 360° in the horizontal axis and from −90° to 90° in the vertical axis [32] (Figs. 14 and 15).

There are various methods to create panoramic images as mentioned by Peleg et al. [47]. Circular projections or rotating cameras are some of the most common and still, as mentioned SVS, it exists a vertical disparity. The vertical disparity is created when the pixel is not founded in the same location of the vertical axis. The points are computed from Eq. (34):

$$\theta = \cos^{-1}\left(\frac{r}{p_x}\right) \quad \phi = \tan^{-1}\left(\frac{p_y}{\sqrt{p_x^2 - r^2}}\right) \tag{34}$$

where a point $(p_x, p_y, 0)$ is projected in the image. The panorama possesses a radius r, and due to the rotational capacities of the cameras, a pair of perspective images is generated from the left and right cameras, C_l and C_r, respectively. Also, an α angle represents the direction of view from the camera [3, 75]:

$$C_{\mathrm{l}} = \begin{bmatrix} \cos\phi \sin\left(\frac{\pi}{2} - \theta - \alpha\right) \\ \sin(\phi) \\ \cos\phi \cos\left(\frac{\pi}{2} - \theta - \alpha\right) \end{bmatrix} \quad C_{\mathrm{r}} = \begin{bmatrix} \cos\phi \sin\left(\theta - \frac{\pi}{2} - \alpha\right) \\ \sin(\phi) \\ \cos\phi \cos\left(\theta - \frac{\pi}{2} - \alpha\right) \end{bmatrix} \tag{35}$$

6 Laser Scanner Systems

The object recognition through laser scanner systems (LSS) is a methodology employing photoelectronic instruments capable of detecting the light emitted by a laser. There exist different types of LSS and a diversity of photoelectronic sensors that detect light at different speeds. Therefore, it is important to note the impact of the methodologies in the precision of the point estimation; the different strategies improve the resolution incrementing the number of detected points during the scanning.

One approach to compute measurements is with the help of triangulation methods. Laser triangulation can be static, where the method requires an adjusted and fixed laser and camera to capture the light. Hence, a dynamic laser triangulation method requires a moving laser until the point of light goes through the selected area or object, and then the triangulation is calculated and the measurement is obtained [53, 54].

The LSS process produces a point cloud representing the object surface, where the points generated must present a low dispersion to reflect the true form of the scanned object. Therefore, despite mainstream methods for registering the measurements as triangulation laser scanner, an error in the estimation of the object's shape is still present. Methodologies as artificial neural networks (ANN) are a helpful solution to LSS, thanks to their capacity to detect and predict patterns also applied in image classification [58, 67] (Fig. 16).

In a navigation system, LSS provides information about the surrounding environment, shapes, size, and depth of objects. As discussed previously, the photoelectronic sensors and laser are key factors in distance measurement, where some lasers are designed to measure large distances. Aerial vehicles as drones require LSS to create maps and determine features for the study of ecologic areas. On the other hand, terrestrial vehicles as cars implement LSS to recognize the road, helping self-driving vehicles avoid collision and measure distances (Fig. 17).

Nevertheless, there exist some lasers in LSS for short distances, applied in the high-precision detection, where they help in differentiation of near objects with detail and provide support for navigation body until they find the desired point. Close-range navigation with laser is applied for microsurgery systems where the level of precision required is high to properly complete the procedure on people. Mobile robots in pipelines realize a mapping of structures to extract features to detect damaged elements such as elbows, T-junctions, or corrosion in the pipeline.

For an LSS working with an INS, the inertial navigation process gets feedback from the last iteration of position and attitude in conjunction with the data coming

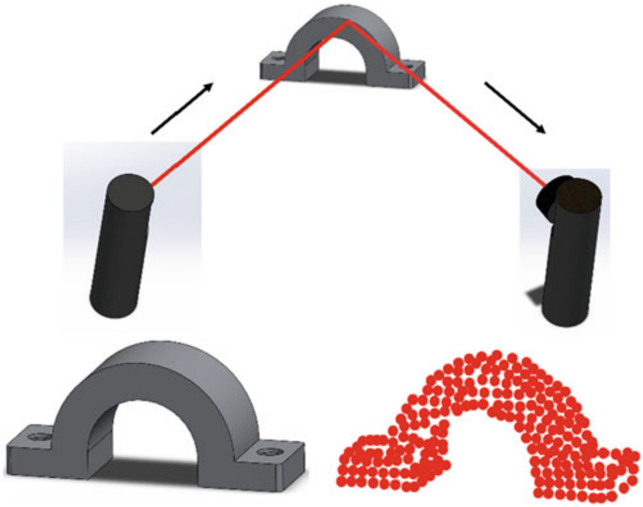


Fig. 16 Laser scanning (upper image) and point cloud (below image)

Fig. 17 Aerial mapping with laser scanners [55]

from the LSS. The LSS system performs a point triangulation to measure the distances and then proceed to apply a transformation of the data in LSS to the body reference frame (Fig. 18).

LSS coordinate reference frame as other reference frames follows the right-hand rule, where Z axis is parallel to the scanning aperture of a laser scanner system and is pointing up, Y axis is the pointing direction of the scanning aperture laser, and X axis is orthogonal to Z axis and Y axis and pointing to the left [68] (Fig. 19).

And to transform the LSS coordinate reference frame to body coordinate reference frame, the system as other references can employ a quaternion transformation matrix or a direction cosine matrix. The present chapter shows a direction cosine matrix required to transform the LSS frame. However, the LSS frame requires

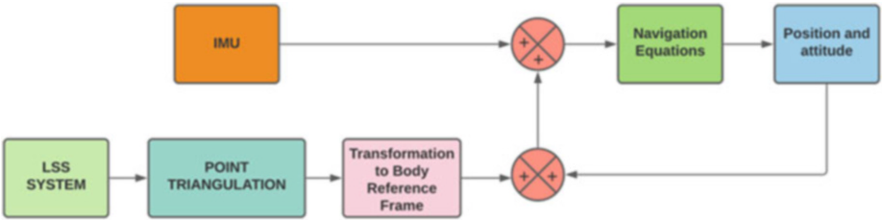


Fig. 18 LSS and INS integration

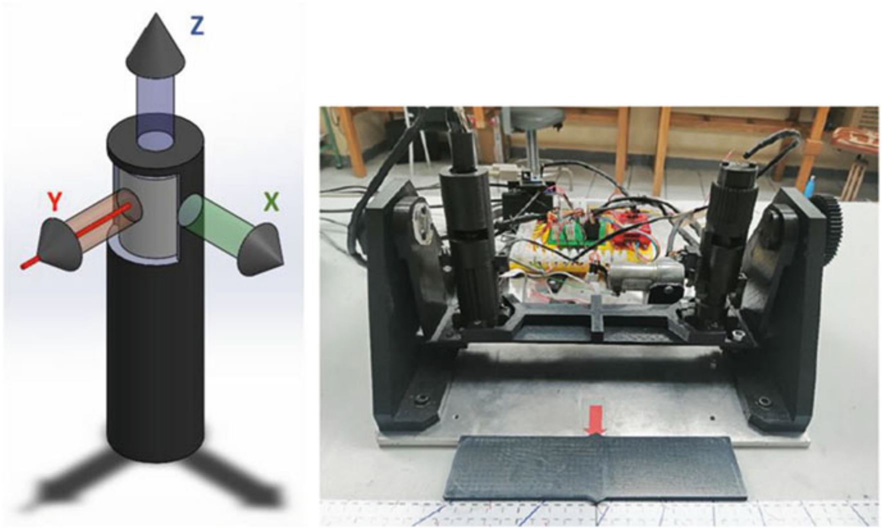


Fig. 19 LSS coordinate reference frame

aligning the laser scanner; the process presents the following equations to properly center the scanner [36, 65]:

$$P^{\mathrm{L}} = \begin{bmatrix} p_i \sin\phi_i \cos\theta_i \\ p_i \sin\phi_i \sin\theta_i \\ p_i \cos\phi_i \end{bmatrix} \tag{36}$$

For Eq. (6), P^{L} is the offset to properly align the laser scanner to the system, ϕ is the mirror angle, θ is the laser scanning axis, and p_i is an offset of the pulse on the distance measurement. The DCM complements the aligning process. As a result, to transform the LSS measurement to body reference frame, the equation requires the rotation in X and Y axis:

$$C_{\mathrm{la}}^{\mathrm{b}} = C_{\mathrm{l}}^{\mathrm{b}} P^{\mathrm{L}} \tag{37}$$

where $C_{\mathrm{la}}^{\mathrm{b}}$ is the DCM of the aligned laser in the body reference frame and $C_{\mathrm{l}}^{\mathrm{b}}$ represents only the rotation from laser frame to body frame. The transformation performs a rotation in X, Y, and Z axis as shown before in Eq. (38):

$$C_{\mathrm{l}}^{\mathrm{b}} = \begin{bmatrix} \cos\psi\cos\theta & \sin\psi\cos\theta & -\sin\theta \\ -\sin\psi\cos\phi + \cos\psi\sin\theta\sin\phi & \cos\psi\cos\phi + \sin\psi\sin\theta\sin\phi & \cos\theta\sin\phi \\ \sin\psi\sin\phi + \cos\psi\sin\theta\cos\phi & -\cos\psi\sin\phi + \sin\psi\sin\theta\cos\phi & \cos\theta\cos\phi \end{bmatrix} \tag{38}$$

7 LIDAR Odometry and Mapping

The motion estimation is a task discussed in different parts of the present chapter. INS has the capability to perform odometry through different means and work in real time. LSS systems are capable of executing the task, and due to their capacity to characterize forms and their measurements, mapping in real-time using LSS is still a popular technology.

On the other hand, light detection and ranging (LIDAR) is a measurement technology based on laser technology. The system principle consists of a transmitter and a receiver; LIDAR measures the time it takes for the laser to travel to a point and go back to the receptor. Therefore, a common practice for the improvement of precision is comparing the measurements to other instruments' data. Consequently, LIDAR technology is used for mapping in topography and exploration [70] (Fig. 20).

The mapping process requires the LIDAR system in movement; as a consequence, there are a series of problems affecting the precision in the measurements. In order to implement a LIDAR in an INS for a mapping process in real time, it is

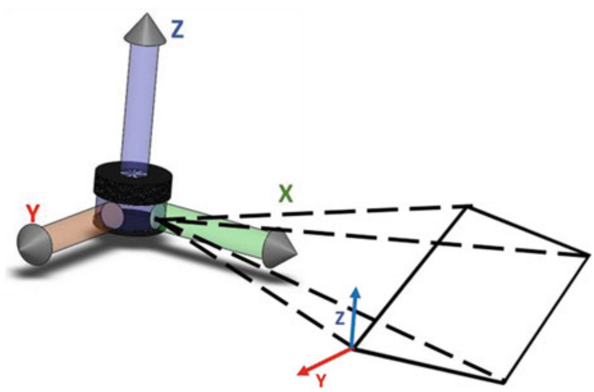
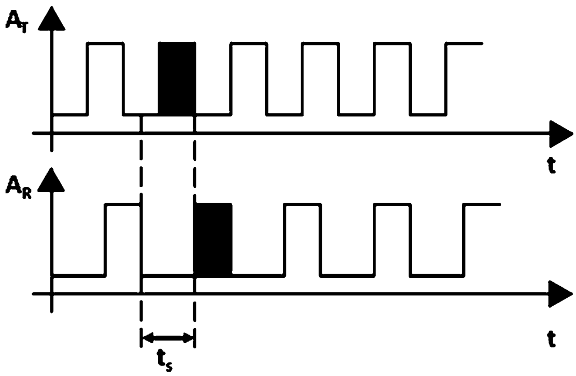


Fig. 20 LIDAR system and FOV

Fig. 21 Traveling time t_s of a laser pulse



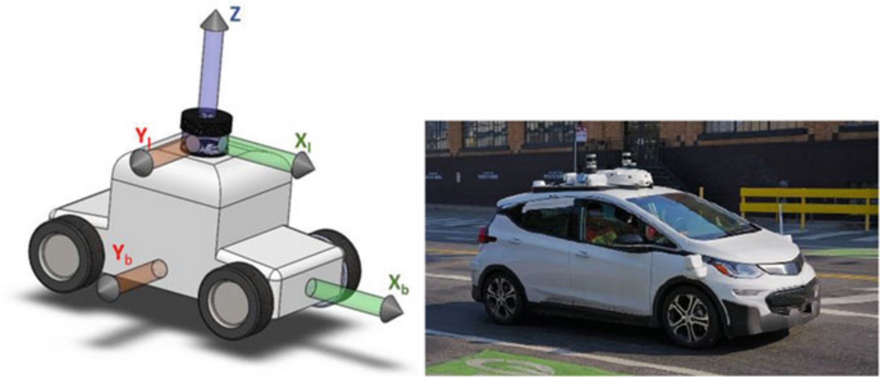


Fig. 22 Body and LIDAR coordinate reference frame and car with a LIDAR [18]

necessary to consider the synchronization of the data received. LIDAR measures the timestep a light pulse takes to reach the target and back to the receiver [44, 79].

The range of LIDAR pulse of light is calculated as follows:

$$R = \frac{1}{2} c \bullet t_s \tag{39}$$

where R is the range or distance between the laser transmitter and the surface object, c is the speed of light, and t_s is the traveling time of the laser pulse (Fig. 21). The amplitudes of the laser pulse in Fig. 22 demonstrate the traveling time t_s [17]. A_T and A_R are the amplitude transmitted and the amplitude received, respectively.

The information proportioned by LIDAR is matched with the inertial instruments as in other INS. Besides the synchronization in time of the data samples, the LIDAR information must be in the body reference [34]. Henceforth, due to LIDAR scanning laser process, the system can share the same coordinate reference frame for X and Y axis. Z axis is pointing in the opposite direction.

After LIDAR initiates the scanning, the data is stored in submaps to form the map of the scanned zone. A variety of methodologies [17, 45, 63] are well known to proceed with the mapping, but in an INS, the information is complemented with inertial references as IMU. Hence, the map created and later stored is an accompaniment for the navigation process; it is still necessary for the technology to properly calculate and process the information coming from the inertial sensor to afterward receive feedback from the mapping process.

Therefore, for the integration of LIDAR with an INS, Su et al. [63] propose to employ the timestep i difference as a reference coordinate, where in a next timestep, j is describing the system trajectory c_{ij} according to the IMU data. Then, it is calculated a component vector of the c_{ij} followed by the body and the current pitch and yaw variation, represented with $d\theta$ and $d\varphi$, respectively. In the particular case of ground vehicles, c is considered as the ground roughness. The component vector possesses a chord length of the trajectory:

$$l_{ij} = c_{ij} \cos(c * d\varphi) \tag{40}$$

Thus, a motion vector p_{ij} is determined with the chord length l_{ij} to obtain the variation between the i and j timesteps:

$$p_{ij} = \begin{bmatrix} c_{ij} \cos(c * d\varphi)\cos(d\theta)\cos(d\varphi) \\ c_{ij} \cos(c * d\varphi)\cos(d\theta)\sin(d\varphi) \\ -c_{ij} \cos(c * d\varphi)\sin(d\theta) \end{bmatrix} \tag{41}$$

8 Surgical Navigation Robots

To perform a surgery, accuracy is a key element for surgeons. It takes years of practice to meticulously realize a complex surgery. Thus, new technologies to perform surgeries with precision are now part of some hospitals and are a helpful tool for inexperienced surgeons. Surgical navigation robots (SNR) are manipulated by trained surgeons and built with inertial sensors as accelerometers or IMU to support the movements and improve the precision.

Surgical navigation robots (SNR) are structures composed of robotic arms, INS system, and complementary sensors to increase the precision. A SNR in conjunction with a LSS allows the systems to perform scans and obtain detailed information of the body part in depth and mapping. Thus, the LSS is capable of realizing a mapping process to create a mesh and to compute finite element calculation [19] (Fig. 23).

LSS is implemented at the end of the link of SNR where the surgical tool is located; in some systems, the laser points directly to the position where it is going to be performed the surgical work. The laser position is corrected through inertial sensors and with the help of cameras where the image is displayed to the surgeon.

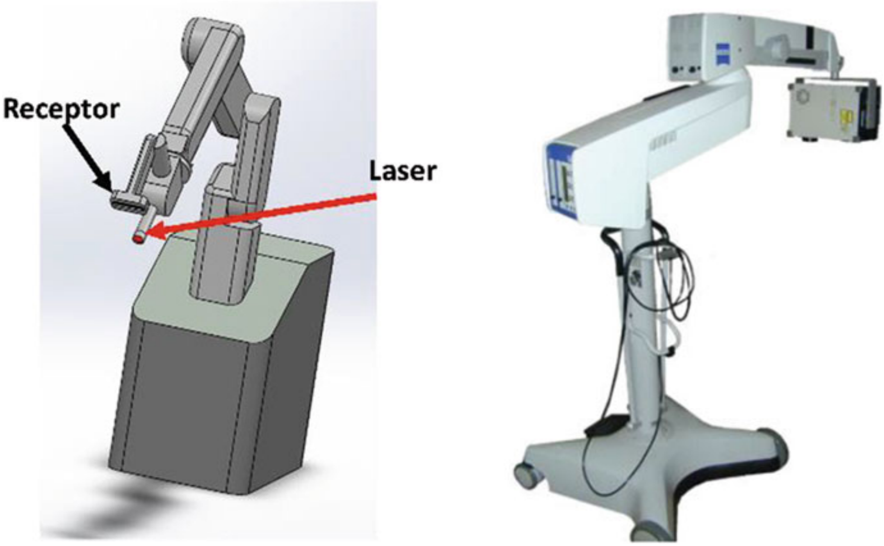


Fig. 23 Laser surgical navigation robot and the Minolta VI-900 class I laser scanner [38]

Robust SNR applies additional optical tracking systems to adjust the laser position [14] [37].

SNR are robots with a variety of configurations depending on the surgery to be performed. Specifications as the number of links for the robotic arm, coupling with other robotic manipulators to improve the precision in the surgery, are elements to consider where the mathematical model is described. For the purposes of the present chapter, the following Eqs. (42, 43, and 44) correspond to the final link of a robotic arm where an LSS could be found. Thus, elements as the body coordinate reference frame and distances for offset laser must be considered.

For a mapping coordinate frame m, Liao et al. [35], Jerbić et al. [27], and Al-Durgham et al. [2] suggest:

$$r_p^m(t) = r_b^m(t) + C_b^m(t)\left\{a_{\mathrm{IMU}/s}^b + C_s^b r_p^s(t)\right\} \tag{42}$$

where ${r_p}^m$, ${r_b}^m$, and ${r_p}^s$ are the positions of the body frame, point p at the end of the arm and a function of the observed range, respectively. ${C_b}^m$ and ${C_s}^b$ are DCM for the rotations; thus, the equations can vary depending on the arm configuration. And $a_{\mathrm{IMU}/s}^b$ is the lever arm offset (if exists) between the laser and the body frame.

Finally, for a validation in the accuracy of measurements Chen et al. [13] propose a pivot P and axis A calibration, where the actual i points collected are compared to the estimated pivotal points $P*$, to obtain a distance error:

$$P_{i\mathrm{err}} = \sqrt{(P_{ix} - P*_{ix})^2 + \left(P_{iy} - P*_{iy}\right)^2 + (P_{iz} - P*_{iz})^2} \tag{43}$$

And for the angle error are included the actual angles A and the calculated angle $A*$:

$$A_{i\mathrm{err}} = \cos^{-1}\left[\frac{A_{ix} - A*_{ix} + A_{iy} - A*_{iy} + A_{iz} - A*_{iz}}{\sqrt{A_{ix}{}^2 + A_{iy}{}^2 + A_{iz}{}^2} \times \sqrt{A*_{ix}{}^2 + A*_{iy}{}^2 + A*_{iz}{}^2}}\right] \tag{44}$$

9 Conclusions

For the integration of a reference to INS, even for similar types of instrumentation, proper identification of the coordinate reference frame and interpretation of the data expressed are required. The data from both SVS and LSS must be transformed to the navigation frame or the measurement frame where it is needed for analysis. All the transformation matrices in this chapter are DCM, but the described methodologies could also be applied for quaternions if the reader is more familiar with them.

When a SVS is incorporated into an INS, it enhances the possibility to attach two or more cameras in a different set of configurations for the system. For a proper interpretation of the data obtained from the image, it is necessary to define where the epipolar line is located with their epipolar points and consider the line of sight and FOV demanded for every camera at the moment to fix them in their location.

Besides, when more cameras are added to the system, a robust calibration is demanded in order to diminish errors in the measurements.

SVS are more common to be implemented with INS or navigation systems; the reason is their low complexity in data interpretation, and they are fast systems. Thus, SVS have the property to determine the objects in the environment even without the involvement of a methodology; the user can simply get the image and correct the trajectory if the system is sending information in real time.

LSS, on the other hand, is more useful for mapping in navigation. The mapping possibility helps the system to recognize a familiar environment, avoid collisions, and improve the navigation. Two of the mentioned LSS own different methodologies to perform measurements during navigation were triangulation and time of flight.

LSS systems could require more time than SVS to perform their scanning and be able to recognize the objects or the structure in front of the body. But a properly scanned object can provide useful information and help the system to perform precise movements as required in medical surgeries.

For both systems, it is necessary to remember that the body could make movements that can distort the camera image and the laser reception, generating measurement errors and increments in the INS drift. The absolute references aid the IMU of the INS, but in other situations, the IMU is helping the vision systems in the navigation, as in the LIDAR case. During the integration of the vision system with

INS, it is desirable to define which system is the one receiving feedback through the navigation process or if both systems are going to work parallel to each other.

References

1. Achtelik M, Bachrach A, He R, et al (2009) Stereo vision and laser odometry for autonomous helicopters in GPS-denied indoor environments. In: Gerhart GR, Gage DW, Shoemaker CM (eds) SPIE Defense, Security, and Sensing. p 733219
2. Al-Durgham K, Lichti DD, Kwak E, Dixon R (2021) Automated Accuracy Assessment of a Mobile Mapping System with Lightweight Laser Scanning and MEMS Sensors. Applied Sciences 11:1007. https://doi.org/10.3390/app11031007
3. Anderson R, Gallup D, Barron JT, et al (2016) JUMP: Virtual reality video. ACM Transactions on Graphics 35:1–13. https://doi.org/10.1145/2980179.2980257
4. Antigny N, Uchiyama H, Servières M, et al (2019) Solving Monocular Visual Odometry Scale Factor with Adaptive Step Length Estimates for Pedestrians Using Handheld Devices. Sensors 19:953. https://doi.org/10.3390/s19040953
5. Bekir, E. (2007). Introduction to modern navigation systems. Introduction to Modern Navigation Systems, 1–235. https://doi.org/10.1142/6481
6. Ben Y, Yang J, Yin D, Li Q (2014) System reset of strapdown INS for pipeline inspection gauge. Ocean Engineering 88:357–365. https://doi.org/10.1016/j.oceaneng.2014.07.004
7. Börjesson, N. (2005). Transformations between Camera Images and Map Coordinates with Applications. Institutionen för systemteknik.
8. Bose A, Bhat KN, Kurian T (2013) Fundamentals of Navigation and Inertial Sensors by Amitava Bose, K.N. Bhat, Thomas Kurian: New Softcover (2013) First edition. | BookVistas. PHI Learning
9. Brooks, S., 2021. https://unsplash.com/photos/jRFJ28gq8ME. [image]. [Accessed 28 March 2021].
10. Castelli, M., 2021. https://unsplash.com/photos/l74VaCZon7I. [image]. [Accessed 28 March 2021].
11. Castro-Toscano MJ, Rodríguez-Quiñonez JC, Hernández-Balbuena D, et al (2018) Obtención de Trayectorias Empleando el Marco Strapdown INS/KF: Propuesta Metodológica. Revista Iberoamericana de Automática e Informática industrial 15:391–403. https://doi.org/10.4995/riai.2018.8660
12. Chauhan T, Héjja-Brichard Y, Cottereau BR (2020) Modelling binocular disparity processing from statistics in natural scenes. Vision Research 176:27–39. https://doi.org/10.1016/j.visres.2020.07.009
13. Chen X, Xu L, Wang H, et al (2017) Development of a surgical navigation system based on 3D Slicer for intraoperative implant placement surgery. Medical Engineering and Physics 41:81–89. https://doi.org/10.1016/j.medengphy.2017.01.005
14. Choi J, Kim J, Hwang JY, et al (2017) A novel smart navigation system for intramedullary nailing in orthopedic surgery. PLOS ONE 12:e0174407. https://doi.org/10.1371/journal.pone.0174407
15. Distante A, Distante C (2020) Paradigms for 3D Vision. In: Handbook of Image Processing and Computer Vision. Springer International Publishing, Cham, pp 315–411
16. Dong Y, Wang D, Zhang L, et al (2020) Tightly Coupled GNSS/INS Integration with Robust Sequential Kalman Filter for Accurate Vehicular Navigation. Sensors 20:561. https://doi.org/10.3390/s20020561
17. Dong P, Chen Q (2017) LiDAR Remote Sensing and Applications. CRC Press
18. Dllu, 2021. A Cruise Automation Chevrolet Bolt. [image] https://commons.wikimedia.org/wiki/File:Cruise_Automation_Bolt_EV_third_generation_in_San_Francisco.jpg [Accessed 1 April 2021].

19. Esslinger D, Rapp P, Knödler L, et al (2019) A novel finite element model–based navigation system–supported workflow for breast tumor excision. Medical and Biological Engineering and Computing 57:1537–1552. https://doi.org/10.1007/s11517-019-01977-0
20. Grewal M, L.~R.~Weill, Andrews A (2007) Global Positioning Systems, Inertial Navigation, and Integration
21. Griesbach D, Baumbach D, Zuev S (2014) Stereo-vision-aided inertial navigation for unknown indoor and outdoor environments. In: 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE, pp 709–716
22. Hanson, A. Andrew J. (2006). *Visualizing quaternions*. 498.
23. He M, Zhu C, Huang Q, et al (2020) A review of monocular visual odometry. Visual Computer 36:1053–1065. https://doi.org/10.1007/s00371-019-01714-6
24. Hofmann-Wellenhof B, Legat K, Wieser M (2011) Navigation: Principles of Positioning and Guidance. Springer Science & Business Media
25. Huang G (2019) Visual-Inertial Navigation: A Concise Review. In: 2019 International Conference on Robotics and Automation (ICRA). pp 9572–9582
26. Jazar, R. N. (2010). *Theory of Applied Robotics Kinematics, Dynamics, and Control (2nd Edition)*.
27. Jerbić B, Švaco M, Chudy D, et al (2020) RONNA G4—Robotic Neuronavigation: A Novel Robotic Navigation Device for Stereotactic Neurosurgery. In: Handbook of Robotic and Image-Guided Surgery. Elsevier, pp 599–625
28. Kelly J, Sukhatme GS (2007) An experimental study of aerial stereo visual odometry. In: IFAC Proceedings Volumes (IFAC-PapersOnline). IFAC Secretariat, pp 197–202
29. Kim H, Lee S (2012) Simultaneous line matching and epipolar geometry estimation based on the intersection context of coplanar line pairs. Pattern Recognition Letters 33:1349–1363. https://doi.org/10.1016/j.patrec.2012.03.014
30. Kuipers, J. B. (2020). *Quaternions and rotation sequences : a primer with applications to orbits, aerospace, and virtual reality*.
31. Kwon H, Park J, Kak AC (2007) A New Approach for Active Stereo Camera Calibration. In: Proceedings 2007 IEEE International Conference on Robotics and Automation. pp 3180–3185
32. Lai PK, Xie S, Lang J, Laqaruere R (2019) Real-time panoramic depth maps from omni-directional stereo images for 6 dof videos in virtual reality. In: 26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 – Proceedings. Institute of Electrical and Electronics Engineers Inc., pp 405–412
33. Li G, Yu L, Fei S (2020) A Binocular MSCKF-Based Visual Inertial Odometry System Using LK Optical Flow. Journal of Intelligent and Robotic Systems: Theory and Applications 100:1179–1194. https://doi.org/10.1007/s10846-020-01222-z
34. Liang S, Cao Z, Guan P, et al (2020) A Novel Sparse Geometric 3-D LiDAR Odometry Approach. IEEE Systems Journal 1–11. https://doi.org/10.1109/jsyst.2020.2995727
35. Liao H, Ishihara H, Tran HH, et al (2010) Precision-guided surgical navigation system using laser guidance and 3D autostereoscopic image overlay. Computerized Medical Imaging and Graphics 34:46–54. https://doi.org/10.1016/j.compmedimag.2009.07.003
36. Liu W, Li Z, Sun S, et al (2019) Improving Positioning Accuracy of the Mobile Laser Scanning in GPS-Denied Environments: An Experimental Case Study. IEEE Sensors Journal 19:10753–10763. https://doi.org/10.1109/JSEN.2019.2929142
37. Luo X, Mori K, Peters TM (2018) Advanced Endoscopic Navigation: Surgical Big Data, Methodology, and Applications. Annual Review of Biomedical Engineering 20:221–251
38. Marmulla R, Hassfeld S, Lüth T, Mühling J (2003) Laser-scan-based navigation in cranio-maxillofacial surgery. Journal of Cranio-Maxillofacial Surgery 31:267–277. https://doi.org/10.1016/S1010-5182(03)00056-8
39. Matthies LH (1989) Dynamic stereo vision. Carnegie Mellon University
40. Mohamed A, Culverhouse P, Cangelosi A, Yang C (2018) Active stereo platform: online epipolar geometry update. EURASIP Journal on Image and Video Processing 2018:54. https://doi.org/10.1186/s13640-018-0292-8
41. Morais, J. P., Georgiev, S., & Sprößig, W. (2014). *Real quaternionic calculus handbook*. 216.

42. Noureldin A, Karamat TB, Georgy J (2013a) Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration. Springer-Verlag, Berlin Heidelberg
43. Noureldin A, Karamat TB, Georgy J (2013b) Basic Navigational Mathematics, Reference Frames and the Earth's Geometry. In: Noureldin A, Karamat TB, Georgy J (eds) Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration. Springer, Berlin, Heidelberg, pp 21–63
44. Palieri M, Morrell B, Thakur A, et al (2021) LOCUS: A Multi-Sensor LIDAR-Centric Solution for High-Precision Odometry and 3D Mapping in Real-Time. IEEE Robotics and Automation Letters 6:421–428. https://doi.org/10.1109/LRA.2020.3044864
45. Pan Y, Han Y, Wang L, Chen J, Meng H, Wang G, Zhang Z, Wang S, (2019) 3D Reconstruction of Ground Crops Based on Airborne LiDAR Technology. In: IFAC-PapersOnLine. 52(24):35–40, ISSN 2405-8963, https://doi.org/10.1016/j.ifacol.2019.12.376
46. Pasqualetto Cassinis, L., Fonod, R., & Gill, E. (2019). Review of the robustness and applicability of monocular pose estimation systems for relative navigation with an uncooperative spacecraft. Progress in Aerospace Sciences, 110, 100548. https://doi.org/10.1016/J.PAEROSCI.2019.05.008
47. Peleg S, Ben-Ezra M, Pritch Y (2001) Omnistereo: Panoramic stereo imaging. IEEE Transactions on Pattern Analysis and Machine Intelligence 23:279–290. https://doi.org/10.1109/34.910880
48. Pesquet-Popescu B, Cagnazzo M, Dufaux F (2014) Chapter 2 – Motion Estimation—A Video Coding Viewpoint. In: Theodoridis S, Chellappa R (eds) Academic Press Library in Signal Processing. Elsevier, pp 27–92
49. Ramezani M, Khoshelham K, Fraser C (2018) Pose estimation by Omnidirectional Visual-Inertial Odometry. Robotics and Autonomous Systems 105:26–37. https://doi.org/10.1016/j.robot.2018.03.007
50. Ramírez-Hernández LR, Rodríguez-Quiñonez JC, Castro-Toscano MJ, et al (2020a) Improve three-dimensional point localization accuracy in stereo vision systems using a novel camera calibration method. International Journal of Advanced Robotic Systems 17:1729881419896717. https://doi.org/10.1177/1729881419896717
51. Ramírez-Hernández LR, Rodríguez-Quiñonez JC, Castro-Toscano MJ, et al (2020b) Stereoscopic Vision Systems in Machine Vision, Models, and Applications. In: Sergiyenko O, Flores-Fuentes W, Mercorelli P (eds) Machine Vision and Navigation. Springer International Publishing, Cham, pp 241–265
52. Rani P, Jangid A, Namboodiri VP, Venkatesh KS (2018) Visual odometry based omnidirectional hyperlapse. In: Communications in Computer and Information Science. Springer Verlag, pp 3–13
53. Real-Moreno O, Castro-Toscano MJ, Cesar Rodríguez-Quiñonez J, et al (2019) Surface Measurement Techniques in Machine Vision: Operation, Applications, and Trends. https://doi.org/10.4018/978-1-5225-5751-7.ch004
54. Real-Moreno O, Castro-Toscano MJ, Rodríguez-Quiñonez JC, et al (2018) Implementing k-nearest neighbor algorithm on scanning aperture for accuracy improvement. In: Proceedings: IECON 2018 – 44th Annual Conference of the IEEE Industrial Electronics Society. Institute of Electrical and Electronics Engineers Inc., pp 3182–3186
55. Roca David, Martínez-Sánchez Joaquín, Lagüela Susana, Arias Pedro, CC BY 4.0 <https://creativecommons.org/licenses/by/4.0>, via Wikimedia Commons. [image] [Accessed 28 March 2021].
56. Roma N, Sousa L (2005) Least squares motion estimation algorithm in the compressed DCT domain for H.26x/MPEG-x video sequences. In: IEEE Conference on Advanced Video and Signal Based Surveillance, 2005. pp 576–581
57. Samama N (2008) A Brief History of Navigation and Positioning. In: Global Positioning: Technologies and Performance. John Wiley & Sons, pp 1–27
58. Sanchez-Castro JJ, Rodriguez-Quinonez JC, Ramirez-Hernandez LR, et al (2020) A Lean Convolutional Neural Network for Vehicle Classification. In: IEEE International Symposium on Industrial Electronics. Institute of Electrical and Electronics Engineers Inc., pp 1365–1369

59. Santoso F, Garratt M, Anavatti S (2017) Visual–Inertial Navigation Systems for Aerial Robotics: Sensor Fusion and Technology. IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING 14:. https://doi.org/10.1109/TASE.2016.2582752
60. Seok H, Lim J (2020) ROVINS: Robust Omnidirectional Visual Inertial Navigation System. IEEE Robotics and Automation Letters 5:6225–6232. https://doi.org/10.1109/LRA.2020.3010457
61. Sharon Y, Lendvay TS, Nisky I (2017) Instrument Orientation-Based Metrics for Surgical Skill Evaluation in Robot-Assisted and Open Needle Driving. arXiv:170909452 [cs]
62. Spong, M. W., & Vidyasagar, M. (Mathukumalli). (2004). Robot dynamics and control.
63. Su Y, Wang T, Shao S, et al (2021) GR-LOAM: LiDAR-based sensor fusion SLAM for ground robots on complex terrain. Robotics and Autonomous Systems 140:103759. https://doi.org/10.1016/j.robot.2021.103759
64. Sun K, Mohta K, Pfrommer B, et al (2018) Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight. IEEE Robotics and Automation Letters 3:965–972. https://doi.org/10.1109/LRA.2018.2793349
65. Talaya J, Alamus R, Bosch E, et al (2004) INTEGRATION OF A TERRESTRIAL LASER SCANNER WITH GPS/IMU ORIENTATION SENSORS
66. Tomažič S, Škrjanc I (2015) Fusion of visual odometry and inertial navigation system on a smartphone. Computers in Industry 74:119–134. https://doi.org/10.1016/j.compind.2015.05.003
67. Trujillo-Hernández G, Rodríguez-Quiñonez JC, Ramírez-Hernández LR, et al (2019) Accuracy Improvement by Artificial Neural Networks in Technical Vision System. In: IECON Proceedings (Industrial Electronics Conference). IEEE Computer Society, pp 5572–5577
68. Van Dang CB, Takahashi K, Phan ATT (2020) Accuracy assessment of 3d point clouds collected by a low cost uav-based laser scanner system. In: Lecture Notes in Civil Engineering. Springer, pp 815–823
69. Veth M (2006) Fusion of Imaging and Inertial Sensors for Navigation. Theses and Dissertations
70. Wandinger U (2006) Introduction to LIDAR. In: LIDAR. Springer-Verlag, pp 1–18
71. Wang, C., Wen, C., Dai, Y., Yu, S., & Liu, M. (2020). Urban 3D modeling with mobile laser scanning: a review. Virtual Reality & Intelligent Hardware, 2(3), 175–212. https://doi.org/10.1016/J.VRIH.2020.05.003
72. Wang D, Li M, Huang X, Zhang X (2021) Optical Autonomous Navigation Technology. Spacecraft Autonomous Navigation Technologies Based on Multi-source Information Fusion 163–210. https://doi.org/10.1007/978-981-15-4879-6_8
73. Wang L, Niu X, Zhang Q, et al (2013) A camera/IMU tightly- coupled navigation algorithm and verification by hybrid simulation. Journal of Harbin Institute of Technology (New Series) 20:84–90. https://doi.org/10.11916/j.issn.1005-9113.2013.06.012
74. Xu G, Zhang Z (1996) Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach. Springer Netherlands
75. Yan J, Kong L, Diao Z, et al (2018) Panoramic stereo imaging system for efficient mosaicking: parallax analyses and system design. Applied Optics 57:396. https://doi.org/10.1364/ao.57.000396
76. Yazdkhasti S, Sasiadek JZ, Ulrich S (2016) Performance enhancement for GPS/INS fusion by using a fuzzy adaptive unscented Kalman filter. In: 2016 21st International Conference on Methods and Models in Automation and Robotics (MMAR). IEEE, pp 1194–1199
77. Yousif K, Bab-Hadiashar A, Hoseinnezhad R (2015) An Overview to Visual Odometry and Visual SLAM: Applications to Mobile Robotics. Intelligent Industrial Systems 1:289–311. https://doi.org/10.1007/s40903-015-0032-7
78. Zhai G, Zhang W, Hu W, Ji Z (2020) Coal Mine Rescue Robots Based on Binocular Vision: A Review of the State of the Art. IEEE Access 8:130561–130575. https://doi.org/10.1109/ACCESS.2020.3009387
79. Zhang J, Singh S (2017) Low-drift and real-time lidar odometry and mapping. Autonomous Robots 41:401–416. https://doi.org/10.1007/s10514-016-9548-2

80. Zhang Y-J (2021) Handbook of Image Engineering. Springer Singapore
81. Cova TJ, Miller HJ, Beard K, Frank AU, Goodchild MF (eds) (2008) Geographic Information Science: 5th International Conference, GIScience 2008, Park City, UT, USA, September 23–26, 2008, Proceedings. Springer-Verlag, Berlin Heidelberg