Research on an Object Positioning Method and Device

Runbo Lu *, Ke Zeng, Jianyu Li *, Huaxi Huang *, Yifan Hu, Qiwei Liu

No. 1 Substation Management Office of Dongguan Power Supply Bureau of Guangdong Power Grid Co., Ltd
Dongguan, China
* Corresponding author: Runbo Lu, Jianyu Li, Huaxi Huang

Abstract: The first position information of each object to be measured in the monitoring area is obtained by a first camera which can position and distance, and then the second camera is placed at the same position of the first camera to obtain the second position information, and the position information of the two is matched by an image matching algorithm to finally realize the object positioning by a monocular camera, and the center of the double antenna and the center of the spatial position sensing device are overlapped together, so that the calculation of coordinate conversion is more simplified, and the generation of errors is also reduced, and the accuracy of orientation positioning is improved.

Keywords: near electrical work; object positioning; electromagnetic interference; multi-eye camera

1. Introduction

In security surveillance, the position of the target in the image can be detected through the camera image, but in practical applications, it is often necessary to know that the position in the image corresponds to the actual geographical position, that is, the latitude and longitude coordinate position. However, because the camera image is in 2D imaging mode, it is difficult to obtain the 3D position information of the pixel coordinate points, so it is difficult to obtain the latitude and longitude coordinate positions of the pixel coordinate points.

At present, the two cameras of binocular cameras are generally used to form a triangular relationship with the monitored object to measure the distance between the object and the camera, so as to locate the monitored object. However, the algorithm of binocular cameras is complex, and the accuracy is related to the distance between binoculars. Only when the distance between binoculars is large can the safety monitoring requirements of large scenes such as substations be guaranteed. Moreover, the price of binocular cameras is higher than that of monocular cameras, which is not conducive to massive promotion. Moreover, it is necessary to obtain the shooting direction of the camera. The shooting direction includes roll angle, pitch angle and yaw angle. The traditional yaw angle acquisition device relies on the geomagnetometer device, but in places with strong electromagnetic field interference, such as substations or power plants, etc., the geomagnetometer is easily interfered, affecting the accuracy and stability of orientation.

This achievement provides an object positioning method and device [1,2], which makes the calculation of coordinate conversion more simplified, and at the same time, it can reduce the generation of errors and improve the accuracy of positioning and orientation.

2. Composition and Principle of Object Positioning Device

The object positioning device, that is, the first camera, refers to a multi-eye camera[3,4], that is, a three-fixed camera capable of positioning, orientation and distance, and the three-fixed camera includes an image recognition unit, an object spatial position sensing unit, a GNSS positioning, a GNSS orientation, a gravity inclination measuring unit (pitch inclination and roll inclination sensor), a pan/tilt and its supporting device, a wireless communication device, a cloud server processing, a speaker acoustic, an object positioning coordinate mapping unit and an image pixel matching unit.

The object positioning device comprises a first position information acquisition module, a second position information acquisition module and an object positioning module to be measured. A first position information acquisition module, configured to acquire first position information of each object to be measured in the monitoring area through a first camera; A second position information acquisition module, configured to determine a camera position corresponding to the first camera, and acquire second position information of each object to be measured in the monitoring area by the second camera based on the camera position, wherein the first camera is a multi-ocular camera and the second camera is a monocular camera; The object to be measured positioning module is configured to position each of the objects to be measured according to the first position information and the second position information through a second camera. The object to be measured positioning module is configured to position each of the objects to be measured according to the first position information and the second position information.

The first position information acquisition module includes a coordinate system construction and three-dimensional coordinate determination unit, a camera attitude and rotation angle acquisition unit, and a first position information determination unit.

A coordinate system constructing and three-dimensional coordinate determining unit configured to construct a positioning coordinate system based on the first camera and the monitoring area, respectively, and determine a camera center three-dimensional coordinate of a center point of the first camera, wherein the positioning coordinate system includes a world three-dimensional coordinate system and a camera coordinate system; A camera attitude and rotation angle acquisition unit configured to acquire a camera attitude and a camera rotation angle corresponding to the first camera; A first position information determining unit configured to determine first position information of each of the object to be measured according to the camera center three-dimensional coordinates, the camera attitude, and the camera rotation angle. The first position information determining unit is configured to determine the first position information of the object to be measured according to the camera center three-dimensional coordinates, the camera posture, and the camera rotation angle.

The camera attitude and rotation angle acquisition unit includes a camera rotation angle acquisition sub-unit, a non-interference situation attitude acquisition sub-unit, and an interference situation attitude acquisition sub-unit.

A camera rotation angle acquisition subunit, configured to acquire a camera rotation angle of the first camera through a gravity inclination sensor, and acquire an electromagnetic interference situation[5,6], wherein the electromagnetic interference situation includes non-severe interference and severe interference; A non-interference situation attitude acquisition sub-unit configured to determine a rotation angle of the first camera with respect to a reference axis as the camera attitude when the electromagnetic interference situation is non-severe interference; An interference situation attitude acquisition sub-unit configured to acquire a camera attitude of the first camera through a dual-antenna sensor when the electromagnetic interference situation is severe interference; An interference situation attitude acquisition sub-unit configured to acquire a dual-carrier phase difference of a satellite signal by the dual-antenna sensor, substitute the dual-carrier phase difference into a first preset algorithm to determine an integer ambiguity, and substitute the integer ambiguity into a second preset algorithm to determine a baseline vector; Substituting the baseline vector into a preset coordinate transformation algorithm to calculate a yaw angle, and taking the yaw angle as the camera pose.

3. Detailed Embodiment of Object Positioning Method

The object positioning method is suitable for the case of movement detection of an object in a monitoring area, and the method can be executed by an object positioning device, which can be implemented in the form of hardware and/or software. The specific implementation of the object positioning method is analyzed in detail as follows:

1. Acquiring first position information of each object to be measured in the monitoring area through a first camera.

The first camera refers to a multi-eye camera, that is, a three-fixed camera that can perform positioning, orientation and distance. Multi-vision technology is a cutting-edge computer vision technology that uses two or more sensors (such as lidar) or cameras (photoreceptors) to obtain multiple perspectives of the scene, thereby realizing the perception and understanding of three-dimensional space. The fixed-range orientation function can also be realized by replacing the multi-eye camera in the present application with a multi-line laser radar. A monitoring area refers to a specific spatial range that is monitored and observed by a specific device or system. The monitoring area can be a room, building, outdoor venue, factory floor or traffic road, etc. The present achievement provides a three-fixed camera, which includes an image recognition unit, a spatial position sensing unit, a pan/tilt, a dual antenna, a multi-axis sensor such as a gravity inclination sensor, and a magnetometer.

An image recognition unit that can collect on-site two-dimensional image data by relying on a camera for identifying object categories, features and contours; The spatial position sensing device can rely on a multi-eye camera or lidar to perform three-dimensional real-time imaging of the object in front, dynamically collect three-dimensional coordinate data of the object on the spot, and map and register it in the originally built three-dimensional point cloud space, using the spatial position sensing device (For example, the center of a multi-eye camera or lidar) as the origin of imaging. Multi-axis sensor, a high-precision attitude and heading reference system using Microelectro Mechanical Systems (MEMS) technology. The sensor combines gravity inclination measurement, a three-axis accelerometer, a gyroscope, a magnetometer, and an onboard processor. The gravity inclination sensor is the pitch and roll inclination sensor (that is, the level sensor), which is used to monitor and control the level of the pan/tilt. It uses gravity sensing technology to avoid interference from high magnetic field places such as substations; The magnetometer is a heading angle sensor, which is used to monitor and control the southeast, northwest and northwest orientation of the pan/tilt. If the accuracy of the magnetometer is not high in the case of electromagnetic interference, it is instead oriented by GNSS. The pan/tilt is used for supporting the spatial position sensing device and can monitor and control the orientation of the spatial position sensing device.

2. Respectively constructing a positioning coordinate system based on the first camera and the monitoring area, and determining the camera center three-dimensional coordinates of the center point of the first camera, wherein the positioning coordinate system includes a world three-dimensional coordinate system and a camera coordinate system.

Constructing a positioning coordinate system based on the first camera and the monitoring area, respectively, and determining the camera center three-dimensional coordinates of the center point of the first camera, comprising: acquiring a preset reference origin according to the monitoring area, and establishing a world three-dimensional coordinate system based on the reference origin; Establishing a camera coordinate system based on a center point of the first camera; Take the world three-dimensional coordinate system and the camera coordinate system as the positioning coordinate system; The dual antenna sensor obtains the camera center position based on the world three-dimensional coordinate system, and substitutes the reference origin and the camera center position into the preset coordinate conversion algorithm to generate the corresponding camera center three-dimensional coordinates in the world three-dimensional coordinate system.

A reference origin needs to be determined in the monitoring area. The reference origin can be a fixed point in the scene, such as a corner or a specific marker. According to the preset reference origin, a world three-dimensional coordinate system can be established for the monitoring area. Can be represented in a Cartesian coordinate system or other suitable coordinate system. The coordinate axis direction of the coordinate system can be defined according to actual needs, usually including three directions: x, y, and z.

For example, if the exact gps coordinates of a place are known, the three-dimensional coordinate system of the world can be established by taking this place as the reference origin, with due north as the z-axis forward, due east as the x-axis forward, and vertical ground upward as the y-axis forward.

Specifically, the controller also establishes a camera coordinate system using the center point of the first camera as an origin. The camera coordinate system can be a coordinate system in which the camera center is the origin, the camera length right direction is the x-axis forward direction, the height upward direction is the y-axis forward direction, and the width forward direction is the z-axis forward direction.

Further, the first camera is equipped with a dual-antenna positioning sensor, and can acquire accurate GPS data of the camera center, compare it with the GPS origin data, calculate and convert it into a distance, and obtain the xz coordinate of the camera center in the reference coordinate system. If the height of the camera center is known, the y coordinate is determined, and then the xyz three-dimensional coordinate of the camera center in the reference coordinate system is obtained. That is, the dual antenna sensor is used to measure the position information of the camera. It can determine the position of the camera in the world coordinate system by receiving signals or other techniques. Finally, the obtained coordinates of the camera center position and the reference origin are substituted into the preset coordinate conversion algorithm. The coordinate transformation algorithm is used to transform the coordinates in the camera coordinate system into the coordinates in the world three-dimensional coordinate system. Through the calculation of coordinate transformation algorithm, the central three-dimensional coordinates of the camera in the world three-dimensional coordinate system are obtained. This coordinate represents the position of the camera relative to the world coordinate system. The preset coordinate conversion algorithm is expressed by the following formula (1):

$$S = 2\,arcsin\sqrt{Sin^2\frac{a}{2} + cos(Lat1) \times cos(Lat2) \times Sin^2\frac{b}{2}} \times 6378.137 \tag{1}$$

Among them, a is the difference in latitude of two points, a = Lat1-Lat2, Lat1 and Lat2 are the latitude of two points, b is the difference in longitude of two points, b = Lung1-Lung2, Lung1 and Lung2 are the longitude of two points, 6378.137 is the radius of the earth, the unit is kilometers; The calculated result unit is kilometers, which is multiplied by 100000 to convert into centimeters. The value of x is the distance between point O (Lung0 Lat0) and point Mx (Lung1 Lat0); The z value is the distance between point O (Lung0 Lat0) and point Mz (Lung0 Lat1).

3. Acquire a camera pose and a camera rotation angle corresponding to the first camera.

Obtaining a camera pose and a camera rotation angle corresponding to the first camera includes: obtaining a camera rotation angle of the first camera through a gravity inclination sensor; Obtaining an electromagnetic interference condition, wherein the electromagnetic interference condition includes non-severe interference and severe interference; When the electromagnetic interference situation is non-severe interference, taking the rotation angle of the first camera with respect to the reference axis as the camera pose; When the electromagnetic interference situation is severe interference, the camera pose of the first camera is obtained by the dual-antenna sensor.

The rotation angle of the first camera may be measured with a gravity inclination sensor. The gravity tilt sensor is able to sense the change in the direction of gravity, thus providing information on the rotation of the camera in space. At the same time, the controller also needs to determine the situation of electromagnetic interference. Electromagnetic interference can be divided into two cases: non-severe interference and severe interference. When the degree of electromagnetic interference is light, that is, non-severe interference, the rotation angle of the first camera with respect to the reference axis may be directly taken as the pose of the camera. However, when the electromagnetic interference is serious, the measurement of gravity inclination sensor may be affected. In this case, the camera pose of the first camera may be acquired by means of the dual antenna sensor. The dual-antenna sensor is able to more accurately determine the position and orientation of the camera by receiving and analyzing electromagnetic wave signals.

In a specific embodiment, the first camera performs three-dimensional shooting and three-dimensional video acquisition of each object to be measured in the monitoring area in the area where the reference point is located. At this time, the three-dimensional coordinates of the center of the camera are (Mx, My, Mz), and at this time, the posture of the camera is from facing due north parallel to due east to a pillar (vertical ground, passing through the center of the camera) It is the pitch degree of axial left rotation (positive value, right-hand principle, the angle included between the camera coordinate X axis and the reference coordinate X axis), that is, rotating in the y axis. In the case of serious electromagnetic interference, it is obtained by dual-antenna satellite orientation technology. Then take the central axis in the height direction of the camera (passing through the center of the camera) as the axis, that is, the x axis is the axial upward rotation roll degree (positive value, right-hand principle, the angle included between the camera coordinate Y axis and the reference coordinate Y axis), which is obtained by the gravity inclination sensor. Since the first camera is not rotated about the z-axis, the value of yaw is 0.

Acquiring a camera attitude of the first camera through the dual antenna sensor includes: acquiring a dual carrier phase difference of the satellite signal through the dual antenna sensor; Substituting a dual carrier phase difference into a first preset algorithm to determine an integer ambiguity and substituting an integer ambiguity into a second preset algorithm to determine a baseline vector; The baseline vector is substituted into the preset coordinate transformation algorithm to calculate the yaw angle, and the yaw angle is taken as the camera pose.

The key steps of dual antenna orientation algorithm are as follows: 1. Establish a baseline vector model: The distance between two antennas is regarded as baseline vector, and the path of satellite signal propagation to the two antennas is regarded as two parallel paths. This baseline vector model is the basis of the relative positioning algorithm. 2. Solve the baseline vector: Use the carrier phase of the satellite signal to solve the double difference, that is, calculate the baseline vector between two antennas by solving the carrier phase difference of the satellite signal. This process needs to measure the carrier phase of the satellite signal, and accurately calculate the ambiguity of the whole circumference of the carrier phase, so as to obtain an accurate baseline vector. 3. Convert coordinates and calculate heading angle: Through the calculation result of baseline vector and the attitude information of carrier, the yaw angle of carrier (first camera) can be calculated. 4. Using the interferometer principle: The essence of the relative positioning algorithm is the interferometer principle, that is, by arranging two antennas at different positions on the carrier surface, the signals received by these two antennas are used to solve the baseline vector, so as to obtain orientation information.

4. Determine the first position information of each object to be measured according to the three-dimensional coordinate of the camera center, the camera posture and the camera rotation angle.

Determining the first position information of each object to be measured according to the three-dimensional coordinates of the camera center, the camera posture and the camera rotation angle comprises: substituting the camera posture and the camera rotation angle into a preset matrix template to construct a rotation matrix; Acquiring a three-dimensional point cloud of each object to be measured based on a camera coordinate system by a first camera; The product of the rotation matrix and the three-dimensional point cloud is determined, and the product is added to the three-dimensional coordinates of the camera center to generate the first position information of each object to be measured.

The acquired camera pose and camera rotation angle can be substituted into a preset matrix template. Matrix templates are usually predefined according to specific coordinate systems and transformation rules. The constructed rotational torque is expressed by the following formula (2):

$$R_zR_yR_x = \begin{bmatrix} \cos(yaw) & -\sin(yaw) & 0 \\ \sin(yaw) & \cos(yaw) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(pitch) & 0 & \sin(pitch) \\ 0 & 1 & 0 \\ -\sin(pitch) & 0 & \cos(pitch) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(roll) & -\sin(roll) \\ 0 & \sin(roll) & \cos(roll) \end{bmatrix} \quad (2)$$

Among them, Rx, Ry, and Rz respectively represent the three-dimensional coordinates of the camera center, and roll, pitch, and yaw respectively represent the rotation degrees of Rx, Ry, and Rz with x, y, and z as axes.

Using the first camera, a three-dimensional point cloud of each object to be measured can be acquired based on the camera coordinate system. The product of the rotation matrix and the three-dimensional point cloud is then calculated. Finally, the product of the rotation matrix and the three-dimensional point cloud is added with the three-dimensional coordinates of the camera center to obtain the first position information of each object to be measured.

In a specific embodiment, the ultimate transformation equation from the camera coordinates to the reference world coordinates can be expressed by the following formula (3):

$$X_{world} = R_zR_yR_x * X_{cam} + T(M_x, M_y, M_z) \quad (3)$$

Among them, Rx, Ry and Rz respectively represent the three-dimensional coordinates of the camera center, Xcam is the three-dimensional coordinates of the point cloud of each object to be measured in the camera coordinate system, and T (Mx, My, Mz) is the coordinates of the first camera center point in the reference world coordinate system.

3.5. A camera position corresponding to the first camera is determined, and second position information of each object to be measured in the monitoring area is obtained by a second camera based on the camera position, wherein the first camera is a multi-ocular camera and the second camera is a monocular camera.

The camera position corresponding to the first camera is determined, and the second camera is used to acquire second position information of each object to be measured in the monitoring area based on the position. The second camera is a monocular camera, which captures images through only one lens. As can be appreciated, monocular cameras are simpler, less expensive, and in some cases may be easier to deploy than multi-ocular cameras.

6. Positioning each object to be measured according to the first position information and the second position information by the second camera.

In a specific embodiment, a "three-fixed" camera capable of positioning, orientation and distance can be used to acquire the geographical position coordinate of each pixel in the image of the monitored area, then the monocular camera is placed at the same position of the "three-fixed" camera, and the "three-fixed" camera is matched with each pixel of the monocular camera through calculation, so that each pixel in the 2D image acquired by the monocular camera can have geographical position coordinate information, Therefore, a monocular camera is used instead of a "three-fixed" camera to monitor the geographic location coordinate information of the object in the area. Finally, the second camera can accurately position each object to be measured according to the first position information and the second position information.

According to the technical scheme of this embodiment, the first position information of each object to be measured in the monitoring area is obtained by a first camera capable of positioning and fixed distance, and then the second camera is placed at the same position of the first camera to obtain the second position information, and the position information of the two is matched by an image matching algorithm, and finally the object positioning is realized by a monocular camera, and the center of the dual antenna and the center of the spatial position sensing device are overlapped together, so that the calculation of coordinate conversion is more simplified, and the generation of errors can be reduced and the accuracy of orientation positioning can be improved.

4. Summary and Outlook

Firstly, the "three-fixed" camera is installed in the area to be monitored, and the geographical position coordinates of each pixel in the image of the monitored area are obtained through the "three-fixed" camera; Then, the monocular camera is placed at the same position of the "three-fixed" camera, and each pixel of the images of the monocular camera and the "three-fixed" camera is corresponded one-to-one by using image matching technology, so as to obtain the geographical position coordinates of each pixel in the monocular camera, so as to monitor the geographical position of the moving object by using the monocular camera. The monocular camera detects moving objects such as operators or cranes and reports them to the cloud server through the communication module. In this way, it can ensure the safety monitoring requirements of large scenes such as substations, power plants and chemical plants.

Acknowledgement

The research was funded by: Project No.: 031900KZ24040083; Project name: 3D Visual Intelligent Safety Management Device for Substation.

References

[1] Lee, C., & Wang, H. (2020). Advanced Object Positioning Algorithms Based on Sensor Fusion. International Journal of Robotics Research, 39(10), 1123 - 1140.

[2] Jiancun Fan; Wei Zou; Xiaoyuan Dou; et al. (2023). High - Precision Object Positioning Using Machine Learning Approaches. Journal of Computer Vision and Image Understanding, 215, 103578.

[3] Aron Schmied; Tobias Fischer; Martin Danelljan; et al. (2022). 3D Modeling and Reconstruction Using Multi - eye Camera Systems. ACM Transactions on Graphics, 41(6), 1 - 15.

[4] Zhang, L., & Chen, Q. (2023). Applications of Multi - eye Cameras in Surveillance and Monitoring. Journal of Visual Communication and Image Representation, 85, 103823.

[5] Thomas, R., & Green, S. (2018). Understanding and Controlling Electromagnetic Interference in Modern Electronics. IEEE Electromagnetic Compatibility Magazine, 7(3), 45 - 58.

[6] Patel, N., & Liu, X. (2021). Electromagnetic Interference Analysis and Design Considerations for Wireless Communication Systems. IEEE Transactions on Electromagnetic Compatibility, 63(2), 421 - 432.