Research on Vision-based Autonomous Navigation Algorithm for RVD between Spacecrafts

*1Wei Sun, 2Kai Liu, 3Long Chen

1, 3 School of Mechano-electronic Engineering, Xidian University, 710071, Xi’an, China,
*1sunweitom@tom.com, 3 poulo214@hotmail.com,
2 Tsinghua University, 710071, Beijing, China, liukai@tom.com

Abstract. In order to solve the autonomous navigation problem of RVD(Rendezvous and Docking) in near distance (<2m) for the on-orbit service of spacecraft, a vision-guided method based on geometry feature of the spacecraft was proposed to measure the relative pose (position and attitude) between two spacecrafts. Firstly, after smoothing the captured image, edge image was obtained by using fast self-adaptive edge detection. We use the line feature to segment the interesting area of the spacecraft and the geometry feature was used to recognize the interesting areas, and the intersection points of the object were gotten. Secondly, the coordinates of these points in the world coordinate were figured out by the proposed fast stereo matching algorithm and 3D reconstruction technical. Based on these coordinates, the pose with respect to the world coordinate was calculated. Finally, lines of the recognized region were extracted and tracked based on Hough transform. In order to verify the effectiveness of the proposed algorithm, a hardware system was established based on high performance DSP. The results of satellite model experiment demonstrate that the relative position errors are less than ±20mm, relative attitude errors are less than ±2°, and measuring speed is up to 8fps which satisfies the precision and speed requirement of the RVD system. The errors in this measurement system were analyzed.

Keywords: Autonomous navigation; RVD; Object recognition; Optical measurement; Binocular stereo vision;

1. Introduction

Many spacecraft have an orbit control subsystem (OCS) which applies forces in order to maintain the orbital parameters within certain bounds. Rendezvous and docking relates to the procedure within which two spacecrafts—the chaser spacecraft and the target spacecraft come close and combine into one integral spacecraft. Guidance and control schemes for the final approach phase require precise information about both the attitude and position parameters of the target spacecraft relative to the chaser.

Robots are main key to future space exploration, engineers and scientists are focusing on building advanced robots that will delve deeper into space. One of their main tasks of space robots is tracking and capturing the target spacecraft and space junk, the key technology of which is identifying and locating the space target. In most cases, the space target is the non-cooperative target which has no characteristic cursors. To the space robots, tracking and capturing these targets is the main task. With two CCD cameras, binocular stereo vision system can simulate the vision system of human, it can realize stereoscopic vision from the projection points in the two CCD cameras which are projected from one point in the space. Robot can complete the scheduled tasks by passing 3D coordinates of the point to robot control system.

A lot of researches have been done in the field of vision-based pose estimation of spacecraft
rendezvous and docking, and a lot of methods and algorithms have been proposed, especially for the final approach. Other researches focusing on the pose estimation also received preferable results, but most of the proposed works focus on pose estimation based upon at least 4 feature points. Some scholars have proved that the relative pose could be estimated through an iterative algorithm from image information retrieved from feature targets fitted in a certain arrangement, but the algorithm lacks both estimation and correction of the measurement error.

As to above problem, an approach to identifying the rectangular plane of the non-cooperative target and extracting its 4 apexes based on stereoscopic vision was proposed in this paper. We propose a novel feature tracking and pose estimation framework for the final approach of RVD from monocular images. Feature targets were tracked and pose information was iteratively solved from the three point problem. In order to test the real-time performance and accuracy of the proposed pose estimation algorithm, a video camera system composed of featured target spacecraft, camera on chaser spacecraft and other image devices is utilized.

The paper is organized as follows: Section 2 describes the basic constraints from observing a single plane and the calibration procedure. We start with a closed-form solution, followed by nonlinear optimization. Radial lens distortion is also modeled. In the next section, the Hough Transform and related problems was described. Section 4 provides the experimental results. Both computer simulation and real data are used to validate the proposed technique.

2. Measurement Principle

2.1 Camera imaging model

The first type of lens effects a perspective projection of the world coordinates into the image, just like the human eye does. This combination of camera and lens is called a pinhole camera model because the perspective projection can also be achieved if a small hole is drilled in a thin planar object and this plane is held parallel in front of another plane (the image plane).

The pinhole model is the most common model of the camera imaging model which is a kind of optical imaging model. Geometric relations of imaging cameras can be shown by Fig.1. Point $O$ is the camera heart, x-axis and y-axis is parallel for X-axis and Y-axis of the image, z-axis is camera optical axis, which is perpendicular for the image plane. The optical axis intersecting with the image plane is $O$, which is the origin of the image coordinates in millimeter units. The camera coordinate system is composed of point $O$, x-axis, y-axis and z-axis. The origin of the image coordinates in pixels units is in the upper left vertex.
2.2 Model and Calibration of binocular stereo vision

To obtain the 3D information of the object, extrinsic and intrinsic parameters of the cameras was needed, which was obtained by the camera calibration. We use a flexible technique to easily calibrate a camera. It only requires the camera to observe a planar pattern shown at a few (at least two) different orientations. Either the camera or the planar pattern can be freely moved and the motion need not be known.

Firstly, as Fig.2, we should note that the points $P$ are given in a world coordinate system (WCS). To make the projection into the image plane possible, they need to be transformed into the camera coordinate system (CCS). The CCS is defined that its $x$ and $y$ axes are parallel to the column and row axes of the image, respectively, and the $z$ axis is perpendicular to the image plane. The imaging location in the image of space arbitrary point $P$ is shown by the pinhole model in the Fig.1; Secondly, through the rotation matrix $R$ and translation vector $T$, the world coordinate system can be transformed into the camera coordinate system; according to the geometric relationships, the camera coordinate system can be transformed into the image coordinates in millimeters; Finally, according to the number of pixels in unit millimeter, the image coordinates system in millimeter units can be transformed to the image coordinates system in pixel units.

Fig.1. Imaging geometric relations figure

Fig.2. Stereo camera system
We are now ready to describe the projection of objects in 3D world coordinates to the 2D image plane and the corresponding camera parameters. The focus length of the camera: \( f \); scale factor: \( s \); the physical size of every pixel in the X-axis and Y-axis direction: \( d_x, d_y \); the scale factor u-axis: \( a_u = f / d_x \); the scale factor v-axis: \( a_v = f / d_y \). The relations are given as follows:

\[
\begin{bmatrix}
  u \\
  v \\
  1
\end{bmatrix} = \begin{bmatrix}
  a_u & 0 & u_0 & 0 \\
  0 & a_v & v_0 & 0 \\
  0 & 0 & 1 & 0
\end{bmatrix} \begin{bmatrix}
  R \\
  t \\
  0'
\end{bmatrix} \begin{bmatrix}
  X_w \\
  Y_w \\
  Z_w
\end{bmatrix} = MM, X_w = MX_w
\]

Matrix \( M \) is a projection matrix; \( M_1 \) is determined by \( a_u, a_v, u_0, v_0 \), which is related to intrinsic parameters of the camera; \( M_2 \) is determined by the location of the camera and the world coordinate system, which is related to extrinsic parameters.

According to the principle, relationship between the world coordinates of space point \( P \) and the image coordinates in pixel units of its projective points can be shown, in which, \( z_{lc}, z_{rc} \) is the \( z \) axis coordinates of point \( P \) in the camera coordinate system respectively, \( M_l, M_r \) are projection matrixes of left camera and right one.

\[
\begin{bmatrix}
  u_l \\
  v_l \\
  1
\end{bmatrix} = M_l \begin{bmatrix}
  x \_w \\
  y \_w \\
  z \_w
\end{bmatrix}, \begin{bmatrix}
  u_r \\
  v_r \\
  1
\end{bmatrix} = M_r \begin{bmatrix}
  x \_w \\
  y \_w \\
  z \_w
\end{bmatrix}
\]

(1)

After combining with two equations and eliminating \( z_{lc}, z_{rc} \), we can get the following matrix:

\[
K \cdot P_w = U
\]

(2)

The world coordinates of space point \( P \) can be solved by the equation (2) with the least square method, which is shown by the equation (3):

\[
P_w = (K^T K)^{-1} K^T U
\]

(3)

To solve equation (3), firstly, we need get the projection matrixes \( M_l, M_r \) by Camera calibration.

As shown is Fig 2, Camera calibration is a necessary step in 3D computer vision in order to extract metric information from 2D images. we can classify those techniques roughly into photogrammetric calibration and self-calibration. We use the technique which only requires the camera to observe a planar pattern shown at a few (at least two) different orientations. The pattern can be printed by a laser printer and attached to a reasonable planar surface (e.g., a hard book cover). Either the camera or the planar pattern can be moved by hand. The motion need not be known. Both computer simulation and real data experiment have been used to test the proposed technique and satisfied results have been obtained. Compared with classical techniques, the proposed technique is considerably more flexible: Anyone can make a calibration pattern by him/herself and the setup is...
very easy. Compared with self-calibration, it gains a considerable degree of robustness.

3. Vision-based Autonomous Navigation Algorithm

Firstly, the counter of the rectangular plane was extracted by using Sobel operator and self adapting threshold. After words, the four edges of the rectangular plane were extracted by using the Hough transform, through which we can compute the coordinates of the four apexes of the rectangle plane. According to the coordinates, we can obtain the coordinates in the right image, as shown in Fig.3. Finally, we can obtain the 3-D information of the target through the 3-D reconstruction according to the coordinates in the right and left image.

![Flow chart of the Independent Relative Pose Measurement Based on Stereo Vision](image)

Fig. 3 Flow chart of the Independent Relative Pose Measurement Based on Stereo Vision

3.1 Adaptive edge detection

An adaptive edge detection method is proposed for the sensitivity to noise and given threshold of existing Sobel algorithms. The basic principle of the Sobel algorithm is used and the direction of the detection is expanded. The proposed algorithm generates adaptive threshold automatically according to the mean value of the 3×3 area pixels around the detecting pixel and the property of people's vision. It not only can keep advantages of existing Sobel algorithm such as parallel process, fast calculation speed and the comparatively thin edge, but also can play a certain part in restraining noise. Because the edge detected by Sobel algorithm is comparatively thick, and the efficiency of the edge-thin algorithm is lower, the edge-thin algorithm is analyzed and improved. The improved algorithm detects the edge of the image where there exists noise, filters the fake edge, then thins the edge of the image, and finally gets the single pixel edge. By comparing the test results of the adaptive edge detection algorithm with those of the existing Sobel algorithm, the adaptive edge detection method can generate automatically adaptive threshold, improves the performance in getting detailed-edge and restraining noise.

3.2 Feature detecting and tracking

The features are invariant to image scale and rotation, and are shown to provide robust matching
across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. The features are highly distinctive, in the sense that a single feature can be correctly matched with high probability against a large database of features from many images.

Solar panel bracket is generally has the obvious line feature. Compared to the point features, line feature has obvious characteristics: small amount of information and small amount of calculation etc. This paper presents a method based on Hough transform for line feature extraction and tracking to find line feature of bracket.

Based on Hough transform, the specific steps of feature extraction and tracking are given as follows:

(1) Calculation of Hough space $A(\rho, \theta)$, and non maxima suppression.

(2) The first frame, line feature extraction based on Hough transform.

After Hough transform, we can extract the line, including the satellite body contour, solar panels and solar array support. From the numerous lines, we select interest line, this is the key problem of autonomous navigation. According to the line features of solar panel bracket (the two waist of the bracket are parallel lines, the distance $\Delta \rho$ of each parallel lines is small and the angle is the identical), we can quickly extract the interested line segment and get $(\rho_j, \theta_j, A)$ $(j = 1,2…N_l, N_l$ is number of interested line segment).

(3) Other frame, feature tracking based on Hough transformation.

The present accumulator has peaks that may have values greater than the specified threshold. The simplest way to find these peaks is to compare the values of the peaks with the threshold value. Then the peaks with the highest values are compared with other peaks in the neighborhood using predefined windows. We can draw straight lines from these peaks.

According to the point and line duality, extraction of the sense line corresponds to the Hough points in space, so the space in the Hough using point tracking to achieve image space of feature tracking.

3.3 Correspondence Points Matching

The essence of stereo matching is to search for the corresponding point of a given point in two different images. The two points are the projections of the same space point. The common image matching methods include matching based on gray-scale image, image-based features and interpretation of the image, or the combination of various matching methods. In this paper, a combination of various methods is promoted, including epipolar constraint, the gray correlation based on regional, the reference difference gradient restraint and the exclusive constraint [9,10].

Matching problem can be reduced to a one dimensional search by using the epipolar geometry of the cameras. An epipolar plane is a plane that contains the point $P$ and the optical centers of the left and the right cameras. The intersections of such a plane with the left and the right images define a pair of epipolar lines. any point on the left image can be found on the corresponding right epipolar line[8]. The problem is thus reduced to a one-dimensional matching problem along each pair of epipolar lines. Let the CCD be symmetry to the XOZ plane, this is to say that the correspondence points are in the corresponding line of the two CCD. A geometric constraint relationship exists on the same objects’ images under the same world coordinate system. In stereo vision, we can use image points matching to recover this relationship, on the contrary, use the relationship to restrict
the candidate matching points, which transforms the scope of the searching for the corresponding points from two-dimensional plane to a one-dimensional line. By this way, the robustness and accuracy of matching have been greatly improved.

Firstly, we solve the epipolar equation on the basis of camera calibration, which contributes to reduce the matching region to an area around a straight line. Secondly, a gray score is chosen as the comparability test to filter the parts of the false matching points. In the end, the last matching points are got by the reference difference gradient restraint. Finally, after applying the epipolar constraint, the most correlated criteria of gray value can be used to filter a large number of false matching points. It is agreed in this paper: determine a feature point from one image as the center of a window, which calls the feature point neighborhood window. The realization of this method is as shown in figure 4.

3.4 Pose for calculating relationship

Assuming there are two coordinate system: World Coordinate System Ow-xwywzw and Object Coordinate System. The feature point Pi in two coordinate coordinates are respectively: Pwi = (xwi, ywi, zwi)T, Pti = (xti, yti, zti)T, the transformation relationship between the two coordinates is shown by the equation (6), where R is rotation matrix T is Translation matrix .

\[ P_{w_i} = R P_{t_i} + T \]  
(6)

Supposing that the three feature points have been already known, two linearly independent columns vector \( n_1 \) and \( n_2 \) can be figured out by any two of the feature points, if we make \( n_3 = n_1 \times n_2 \), these three columns are linearly independent and it meets the following equation:

\[ n_w = R \cdot n_t \]  
(7)

Among which ,

\[ n_w = \begin{bmatrix} n_{w_1} \\ n_{w_2} \\ n_{w_3} \end{bmatrix}, n_t = \begin{bmatrix} n_{t_1} \\ n_{t_2} \\ n_{t_3} \end{bmatrix} \]

\( n_{w_1}, n_{t_1} \) are coordinates In the World Coordinate System and Object Coordinate System. According to the equation (8), rotation matrix R, Translation matrix T can be acquired.

\[ \begin{bmatrix} R = n_n \cdot n_{n_1}^{-1} \\ T = P_{w_i} - R \cdot P_{t_i} \end{bmatrix} \]  
(8)

we assume that the order from Catch point coordinate system to World Coordinate System is \( z \rightarrow x \rightarrow y \), matrix R is shown by equation (9):

\[ R(\psi, \phi, \theta) = R_3(\theta)R_2(\phi)R_1(\psi) \]

\[ \begin{bmatrix} C\theta C\psi - S\psi S\theta \phi & C\theta S\psi + S\psi C\theta \phi & -C\phi S\psi \\ -C\theta S\phi & C\theta C\phi & S\phi \\ S\phi C\theta + S\psi C\phi & S\phi S\theta - S\psi C\phi & C\phi C\theta \end{bmatrix} \]  
(9)
In which, C, S stands for cosine and sine operations respectively.

According to the values of the R matrix combining with equation (9), the $\varphi$, $\theta$, $\psi$ that are rotated counter clockwise around x, y, z axis are figured out. The method is shown by equation (10).

\[
\begin{align*}
\varphi &= \arcsin \left[ R(2,3) \right] \\
\theta &= \arctan \left( \frac{R(1,3)}{R(3,3)} \right) \\
\psi &= \arctan \left( \frac{R(2,1)}{R(2,2)} \right)
\end{align*}
\]

(10)

4. Experimental Results

The system consists of image acquisition and poses measurement is shown in Fig.5. We use the DSP TMS320DM6467 as the real time embedded platform, the solar panel bracket as the recognition and measurement of the goal.

![Fig. 5 Constitution of measuring system](Image)

The specific structure of image acquisition is shown in figure 6. The left and right camera and two bar-shaped light sources are mounted on the front panel of a robot, the baseline distance of two cameras is L, the image resolution is 704x 576. Light source reduces the variation of external light influence on measurement, and enhances the recognition of regional contrast, improves the image processing algorithm stability. Light source is red LED light source, black and white CCD chips on red (lambda =660nm) is the most sensitive. Compared with traditional fluorescent lamp light source, LED light source has the advantages of long service life, response speed etc..

![Fig. 6 Configuration of the image acquisition](Image)

We apply the above approach to the three dimensional reconstruction. In this system, the image of the surface layer can be obtained by using double-CCD cameras. According to the coarse to fine matching algorithm, the zero-crossing points are matched, then using mathematics model of binocular stereo vision, we can reconstruct 3D configuration of the surface in the coordinate system.
Fig. 4 shows the measurement of the object based on this method. From Fig. 5, we can see the measurement error is reduced to 2~3cm when the real distance is 150~30cm.

Generally, the size of the experiment model used to serve the target satellite is 500mm×500mm×500mm, the satellite model image collected by cameras is shown in figure 7.

Fig. 7 Binocular image of the proposed system

Coordinate system is defined as shown in figure 8. Graph (a) stands for the World Coordinate System. The World Coordinate System zwOwyw is defined on the robot front panel, and the positive direction of Owzw is the direction from right camera to left camera, point Ow locates in the right camera. The graph (b) is the capture point coordinate system; the zcapOcapycap definition of capture point coordinate system is on the Solar panel stents. Ocap point locates in the midpoint of the line l1. The World Coordinate System and Capture Coordinate System are all right hand coordinate system.

In the experiment, line l1 and l2 are interested line, P1, P2, P3, P4 are feature points. According to the feature extraction and the extraction of the line l1 and l2, by l1 and l2, the benchmark image coordinates of P1, P2, P3, P4 can be obtained. The 3D coordinates of point Ot in the World Coordinate System will be figured out by the stereo matching and 3D reconstruction method according to the algorithm presented before. We make

\[ u_1 = O_1P_1, \quad u_2 = O_1P_2, \]

and the pose of the two coordinate system will be solved by the pose for calculating relationship method introduced.

In the experiment the serving satellite model move upward along x and z, making variable motion along x axis and sinusoidal motion along z axis. Fig. 9 shows the video sequence of the serving satellite. From the sequence we can see the changing perspective of its solar panels stand from global to local. Fig. 10 is the result of the self-adaptive video frequency sequence edge detection, by the test result Fig. 11 we can see that through reasonable light design we can ensure that the first frame of image will get rid of the influence of the Solar Array after self-adaptive edge detection, which can reduce the computation load of feature extraction. Linear feature extraction will then be switched into linear feature tracing. The uniqueness of feature guaranteed that the tracing procedure will not be influenced by the solar array.
Fig. 9 Video sequence of the serviced satellite

Fig. 10 Result of self-adaptive edge detection

Fig. 11 Result of points matching

Fig. 12 and fig. 13 demonstrates the measure data of x and z axis. Fig. 14 shows the mean deviation of their relative position. From the results of Experiments, we can see that the relative deviation is less than ±20mm and ±2 degrees, which can satisfy the requirement for measure accuracy.

In the coordinate system, since the xw axis is parallel with the xcap axis, the deviation between φ, θ is 0 degree.
Fig. 13 data of z-axis

Fig. 14 errors of relative pose

5. Conclusion

In this system, High Brightness Strip light sources were used to weaken the influence of light on measurement and make the feature of the object more prominent. Then we analyze the errors in this measure system and get the factors which impact the measurement accuracy. We use effective data processing methods to reduce errors and improve measurement accuracy. The simulation results show that the method can reach the need of measurement accuracy of the relative position and attitude for spacecraft; it is more suitable in on-board real-time computation.

The beneficial effects of the binocular system based stereovision measurement developed by this paper are mainly manifested in[11-13]: 1) Capturing the 3D video images in real time, and the tracked monitoring objects will be not lost. 2) Providing a complete theoretical system and model for the realization of real time tracking of fast-moving targets in large space. 3) We can realize rapid point to point match, which gives a great convenience for the following 3D image processing. 4) No longer need the complicated calibration work, feature extraction is very convenient, and rapid 3D image matching can be achieved. 5) Adopts a kind of uniform coordinates at the image gathering, 3D matching and 3D image reconstruction, to achieve 3D image reconstruction and 3D objects measuring much more easily through geometric calculation.

The system designed and developed in this paper can be widely used in industrial inspection, robot vision, games and other areas. Limitations of the work are obviously, the spacecraft we discussed in this paper is in the shape of regular geometric structure, The next step, in order to obtain these objectives more robust acquisition algorithm, we will continue to research spacecraft identification method.
6. Acknowledgment

This project is supported by the “Fundamental Research Funds For the central Universities” (No.K50511040008), AND Natural Science Foundation of China (No. 60802077, NO.61003196)

References