Evaluating 3D position and velocity of subject in parabolic flight experiment by use of the binocular stereo vision measurement

Tao Jin (金 涛)[1], Hongzhi Jia (贾宏志)[1], Wenmei Hou (侯文玫)[1], Ryo Yamamoto[2], Norihiro Nagai[2], Yusaku Fujii[2], Koichi Maru[2], Naoya Ohta[3], and Kazuhito Shimada[4]

[1]*Shanghai Key Laboratory of Modern Optical System, School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China*

[2]*Department of Electronic Engineering, Faculty of Engineering, Gunma University, Gunma 376-8515, Japan*

[3]*Department of Computer Science, Faculty of Engineering, Gunma University, Gunma 376-8515, Japan*

[4]*Medical Operations, Japan Aerospace Exploration Agency (JAXA), JAXA Houston Office, Houston 77058, USA*

[*]*E-mail: j.billow@yahoo.com*

Received September 25, 2009

A practical method for evaluating the three-dimensional (3D) position and velocity of a moving object used in the parabolic flight experiment is developed by using the binocular stereo vision measurement theory. The camera calibration mathematic model without considering the lens distortion is introduced. The direct linear transformation (DLT) algorithm is improved to accomplish the camera calibration. The camera calibration result and optimization algorithm are used to calculate the object's world coordinate from image coordinate. The 3D position and the velocity of the moving object are obtained. The standard uncertainty in estimating the velocity is 0.0024 m/s, which corresponds to 1% level of the velocity of the object in the experiment. The results show that this method is very useful for the parabolic flight experiments.

OCIS codes: 100.2550, 150.1488, 150.6910.

doi: 10.3788/COL20100806.0601.

The measurement of the binocular stereo vision is one of the important branches in computer vision systems. Because it is similar to the human visual system and has a high accuracy and efficiency, it is widely used in the industrial inspection, object recognition, workpiece positioning, automatically guided robot, and so on[1]. In recent years, many scholars have extensively researched the binocular stereo vision measurement, and almost all of them focus on how to build the camera calibration mathematical models and the target feature point matching algorithm on vision measuring systems. Canadian Space Agency (CSA) used "space vision system" to calculate the location and orientation of a space payload, and presented this information to the operator in the form of both graphical and textual cues. Stirling *et al.* used a camera to measure the kinetics and kinematics for translation motion in microgravity during the parabolic flight[2]. They used four cameras to collect the data of the moving subject when the subject moved toward a goal. However, the method they used is different from the one we will propose and they used many more hardwares than us. As we all know, under the microgravity condition, mass cannot be measured by load cell or balance. The measurement of the body mass is a medical requirement of the International Space Station (ISS) to maintain their health. Its ideal precision range is ± 0.2 kg while the measurement range is from 40 to 110 kg. Current body mass measurement apparatus on the ISS is large and complex. Thus, the National Aeronautics and Space Administration (NASA) of USA and the Federal Space Agency of Russia have developed only a few instruments for astronaut body mass measurement. Their instruments are bulky, not easily deployed in a spacecraft cabin, and only moderately accurate. To overcome these problems, Fujii *et al.* proposed several designs to measure the body mass of astronauts[3−7]. The "space scale" was proposed, using a rubber cord as a source of force. In the space scale, the velocity or acceleration of the object is measured by using optical interferometer. The subject mass, m, can be calculated from the acceleration a and the force F, as $F = ma$. This method can effectively reduce the volume and complexity of the body mass equipment. Fujii *et al.* conducted the parabolic trajectory flight tests for demonstrating the prototype of the space scale[4,6,7]. The conditions of the parabolic flight in the business jet airplane are typically not optimal due to the residual acceleration field, vibration, and aerodynamic noise inside the cabin. For this reason, it is desirable to estimate the three-dimensional (3D) position and attitude of the object.

In this letter, a practical system for evaluating the 3D position of a human subject is described in the parabolic flight experiment. The equipments we used are only two charge-coupled device (CCD) cameras which are standard equipment for the airplane and are very easy to set. The two CCD cameras are fixed on the wall of the airplane used in experiment and must have crossing visual field. Just like human eyes, two CCD cameras simultaneously record the movement of the subject. According to the two video images from the CCD cameras, the binocular stereo vision measurement technology can figure out the objective movement (including movement distance, etc.). The principles are as follows. Firstly, getting the relation between the image pixel coordinate and the world coordinate according to the coordinate transformation; secondly, using the camera calibration algorithm to calculate the internal and external parameters of the two

cameras; finally, calculating target point's world coordinate combined the results of the first and second steps. The velocity of the object can be calculated by the world coordinate. Our system is very easy to use. It does not need object moving straightly like using Doppler interferometer. There are only two steps needed to do. One step is to fix the two cameras. The other step is to calculate the 3D position and velocity of the subject. All calculation will be automatically completed by computers. In this letter, the data of movement change and velocity are given using our system in the experiment. The velocity we used is compared with that measured by the Doppler interferometer. The comparing result shows that evaluating the 3D position of a subject in parabolic flight experiment using our system is possible and effective. The final target of this study is to develop the 3D position and attitude estimation system using the CCD camera in parabolic flight experiment. This will contribute to the space research and development.

Camera calibration in the situation of the 3D machine vision is the process of determining the internal camera geometric and optical characteristics (intrinsic parameters) and/or the 3D position and the orientation of the camera frame relative to a certain world coordinate system (extrinsic parameters). It means inferring 3D information from two-dimensional (2D) computer image. One of the 3D information is about the location of the object, target, or surface feature. For simplicity, if the object is a point feature, the camera calibration provides a way of determining a ray in 3D space on which the object point must lie, given the computer image coordinates[8−10]. With two cameras, the position of the object point can be determined by intersecting the two rays. Both intrinsic and extrinsic camera parameters need to be calibrated. The relationship among the camera model, world space model, and pinhole imaging model is shown in Fig. 1. Here, $o_{\rm c}$ is the camera's optical center, $o_{\rm c}x_{\rm c}y_{\rm c}z_{\rm c}$ is the camera coordinate system, $o_{\rm w}x_{\rm w}y_{\rm w}z_{\rm w}$ is the world space coordinate system, ouv is the image plane coordinate system, $P(x_{\rm w}, y_{\rm w}, z_{\rm w})$ is one point in the world space system, and $P(u_i, v_i)$ is the projection of point $P(x_{\rm w}, y_{\rm w}, z_{\rm w})$ in the image plane. From Fig. 1, we can see that the image plane system's axes u and v are parallel to the camera system's axes $x_{\rm c}$ and $y_{\rm c}$, respectively. Based on the image model, the relationship between the imaging points $P(x_{\rm w}, y_{\rm w}, z_{\rm w})$ and $P(u_i, v_i)$ is[8]

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{\rm w} \\ y_{\rm w} \\ z_{\rm w} \\ 1 \end{bmatrix}, \quad (1)$$

where, u and v are the pixel coordinates, f_x, f_y, u_0, and v_0 are the internal parameters of the camera used in the experiment, r_{11}–r_{33} are the parameters of the camera-related world coordinate system rotation values, satisfying

$$\begin{cases} r_{11}^2 + r_{12}^2 + r_{13}^2 = 1 \\ r_{21}^2 + r_{22}^2 + r_{23}^2 = 1 \\ r_{31}^2 + r_{32}^2 + r_{33}^2 = 1 \end{cases}, \quad (2)$$

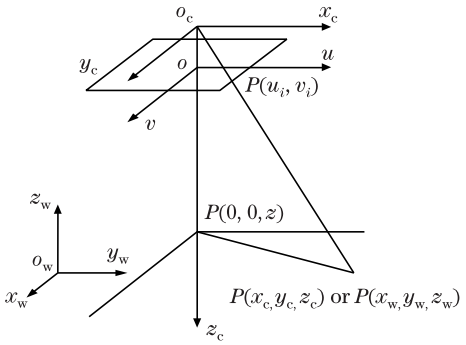

Fig. 1. Imaging camera model.

and T_x, T_y, T_z are the external parameters of the offset of the camera coordinate system relative to the world coordinate system. We set

$$\boldsymbol{M}_{3\times4} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix}. \quad (3)$$

Equation (1) can be unfolded with $\boldsymbol{M}_{3\times4}$ as

$$\begin{cases} \lambda u = m_{11}x_w + m_{12}y_w + m_{13}z_w + m_{14} \\ \lambda v = m_{21}x_w + m_{22}y_w + m_{23}z_w + m_{24} \\ \lambda = m_{31}x_w + m_{32}y_w + m_{33}z_w + m_{34} \end{cases}, \quad (4)$$

where λ is a non-zero constant. To eliminate λ, we can derive

$$\begin{cases} m_{34}u = m_{11}x_w + m_{12}y_w + m_{13}z_w + m_{14} - um_{31}x_w - um_{32}y_w - um_{33}z_w \\ m_{34}v = m_{21}x_w + m_{22}y_w + m_{23}z_w + m_{24} - vm_{31}x_w - vm_{32}y_w - vm_{33}z_w \end{cases}. \quad (5)$$

From it, we can find that one camera has about 12 parameters. If we know 6 control points which are not in the same plane, there will be 12 equations with 12 parameters. The internal and external parameters can be calculated.

Binocular stereo vision systems usually consist of the structure and properties of two identical components of the CCD image sensors. Because two CCDs can shoot the same points at the same time, the CCD's positions are usually symmetrical in front of points[11]. From the camera model and Eqs. (2) and (5), the following superset of equations can be obtained:

$$\begin{bmatrix} u - f_xT_x - u_0T_z \\ v - f_yT_y - v_0T_z \\ u' - f_x'T_x' - u_0'T_z' \\ v' - f_y'T_y' - v_0'T_z' \end{bmatrix} = \begin{bmatrix} f_xr_1 + u_0r_7 & f_xr_2 + u_0r_8 & f_xr_3 + u_0r_9 \\ f_yr_4 + v_0r_7 & f_yr_5 + v_0r_8 & f_yr_6 + v_0r_9 \\ f_x'r_1' + u_0'r_7' & f_x'r_2' + u_0'r_8' & f_x'r_3' + u_0'r_9' \\ f_y'r_4' + v_0'r_7' & f_y'r_5' + v_0'r_8' & f_y'r_6' + v_0'r_9' \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix}, \quad (6)$$

where f', r', and T' are the left camera parameters, f, r, and T are the right ones. These parameters can be

worked out from the camera calibration. (u, v) and (u', v') are the projections of the same point $P(x_w, y_w, z_w)$ in two camera frames. The same point can be fixed according to the shooting rays received by the two cameras. Through the above method, the point P coordinates of the 3D position can be obtained by using least squares method[12].

In the experiment, two cameras need to be calibrated. The right side camera has 6 control points, so we adopt direct linear transform (DLT)[13] algorithm to figure out all the parameters of this camera. But to the data from the camera on the left, as it is with only 5 control points, we cannot apply DLT algorithm to solve all the parameters according to Eq. (5). Many literatures have discussed the details on the 3 or 4 control points[14−16]. They focus on discussing the necessary and sufficient condition of equations and the number of solutions. But for the perspective-five-point (P5P) problem (5 control points), there is less discussion. Some algorithms' robustness per P5P problem is not sufficiently reasonable[17−19]. In this letter, we consider both P5P algorithm and parameters supplied by the CCD manufacturer. Figure 2 shows the example of the frame image of the CCD camera on the parabolic flight test. The object point's pixels must be found out in the two cameras frame images. For extracting the object point from video frame image, motion estimation and advanced Hough transformation are introduced. Motion estimation can extract the moving body and smooth away other interferences. Then the Hough transformation can find out the object point, as shown in Fig. 2. But only using the Hough transformation cannot locate the point precisely. So combining the Hough transformation with the maximum information entropy can extract the object point. The detail of image processing is presented in Ref. [20].

The method process of binocular stereo vision measurement is as follows:

1) using two cameras to record a moving object; 2) decoding of video to frame images; 3) extraction of the control point and the pixel coordinates of target points; 4) using control points and camera calibration algorithms to calculate the transformation matrix of the image coordinate system and world coordinate system; 5) using the binocular stereo vision system mathematical model to solve the moving object's physical coordinates in the world space system.

The microgravity experiment was carries out by the parabolic flight using a jet airplane for evaluating the prototype efficiency of the "space scale". Each experiment was completed during 0-G state within ∼20 s. Before a measurement started, an experimenter stabilized the length before the experimenter released the human subject. The velocity of the body movement was measured with an optical Doppler interferometer. The setup for the "space scale" has been reported in Refs. [6,7].

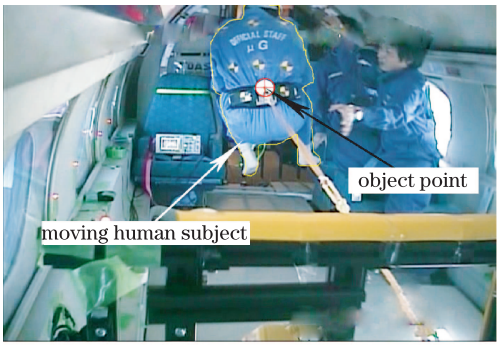

Fig. 2. Example of the frame image of camera.

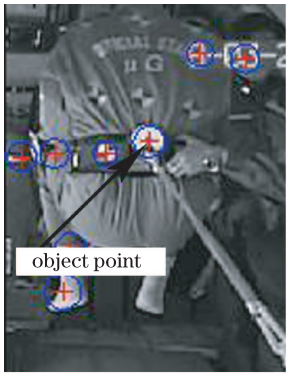

Fig. 3. Results of motion estimation and advanced Hough transformation.

During the experiment, images were recorded on a parabolic trajectory airplane operated by the Diamond Air Service, Inc. (DAS) at Nagoya, Japan. When the gravity became close to zero, the object body moved along the direction of the spring which connected a transducer. The body was photographed by two cameras at the same time. In this experiment, we collected 4.97185-s video and the inter-frame interval was 0.0333 s. We dealt with 51 frame images within 1.7 s. The control points and cameras distance were measured previously. The cameras were set at auto-shutter speed, manual-gain, and auto white balance. The camera controller is CC421 and the camera lens is T1675F.

The results of motion estimation and advanced Hough transformation are shown in Fig. 3. The motion estimation can effectively narrow down the detection region. Even in this kind of region, using the Hough transformation cannot find out the point well. But together with the maximum information entropy method, as shown in Fig. 4, the problem can be solved with reasonable precision. From Fig. 4, we can see that the highest point's coordination is the object point's pixel coordination, and the vertical direction is the total gray value of circle defined by the Hough transformation.

Figure 5(a) is the variation of the motion in the x direction and Fig. 5(b) is the change in the y direction according to the world coordinate. The distance changes are about 110 and 100 mm in the x and y directions, respectively.

Figure 6 is the time versus distance along the z direction in the world coordinate system. Comparing Fig. 5 with Fig. 6, we can find that the object translates 440 mm in the z direction, which is larger than those in x and y directions. It is shown that before 0 s the object point almost does not move in the z direction. But between 0 and 1.6 s the distance changes very quickly, it means that the object point moves very fast during this time. The phenomenon corresponds to the actual situation. So from Figs. 5 and 6, we can estimate the object trajectory and the 3D position of the object. The object

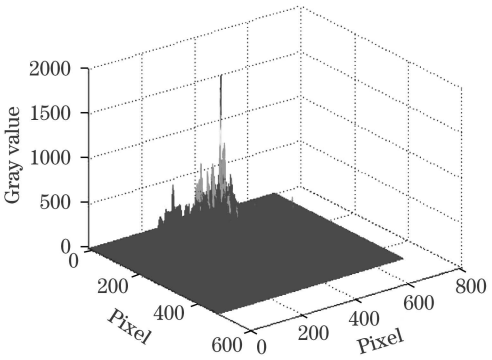

Fig. 4. Gray value accumulation results for the circle detected in Fig. 3.

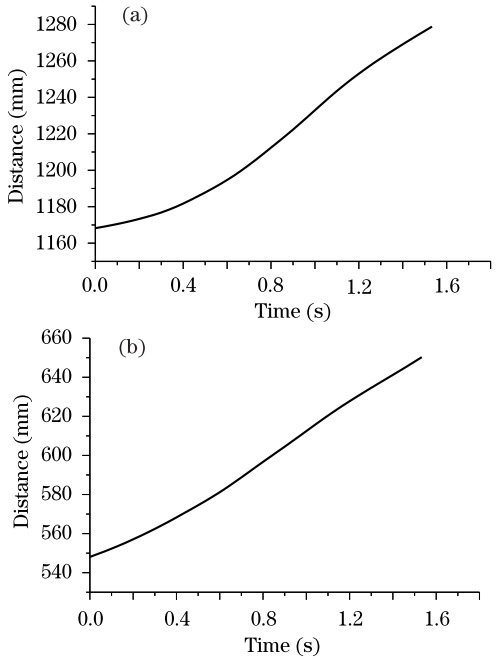

Fig. 5. Object point trajectories in the (a) x and (b) y directions on the world system coordination.

velocity can be calculated by the displacement and time, as shown in Figs. 5 and 6.

The velocity versus the time is shown in Fig. 7. Triangle data are measured by the binocular stereo vision, and diamond data are measured by the Doppler interferometer. For the data measured by the Doppler interferometer, only the diamond data with sufficient accuracy are plotted. The velocities measured by the binocular stereo vision are defined as the component of the direction of the laser beam used for the Doppler interferometer. The time range of the diamond data is from 0.38 to 0.59 s and sampled 248 points. The time of the triangle data is from 0 to 1.53 s and sampled 46 points.

The relative standard uncertainty of optical interferometer is 0.0003 m/s. The root mean square (RMS) value of error is 0.0022 m/s. Therefore, the standard uncertainty in measuring the velocity is estimated to be 0.0024 m/s. It proves that the 3D position and the velocity of subject in the parabolic flight experiment can be estimated with a sufficiently small uncertainty by using the developed method.

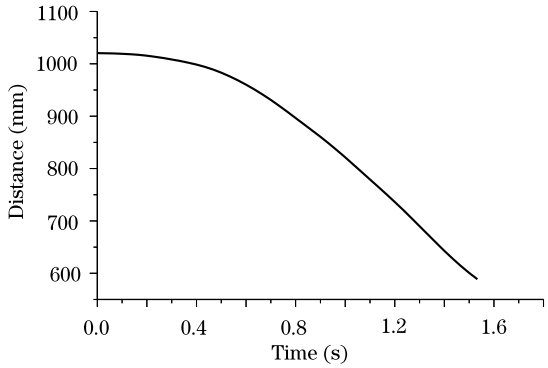

Fig. 6. Object point's trajectory in the z direction.

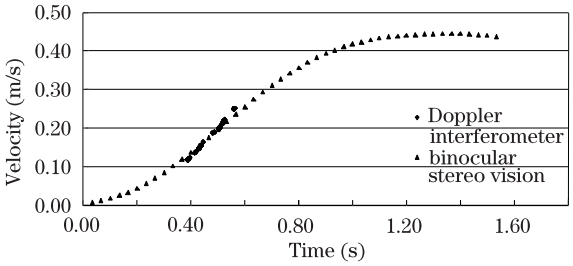

Fig. 7. Velocity data measured by Doppler interferometer and binocular stereo vision.

The result will be better if the following elements are taken into account. First of all, the camera calibration model is an ideal pinhole model which does not consider the lens distortion and the sample error. Secondly the control points should increase. If the control points are enough, the calibration result can be optimized by using Levenberg-Marquardt (LM) algorithm or other camera calibration methods.

In conclusion, a 3D position and attitude estimation system using the CCD camera in the parabolic flight experiment is developed. In the method, only the CCD cameras are used. Up to now, the images taken by those cameras are just used to be looked by eyes, and no further image processing has been done. Our system is easy to use, and no additional hardware is necessary. The developed method can be an essential part of the parabolic flight experiment system. The efficiency of the developed system (method) is demonstrated by the parabolic flight experiment using DAS Gulfstream II airplane. This will be a contribution to the space research and development.

The parabolic flight experiment was conducted as a part of the "Ground-Based Research Announcement for Space Utilization" promoted by the Japan Space Forum (JSF) with the cooperation of DAS.

This work was supported by the Innovation Fund Project for Graduate Student of Shanghai (No. JWCXSL0902), the Shanghai Leading Academic Discipline Project (No. S30502), and the Program from Shanghai Committee of Science & Technology (No. 08DZ2272800).

References

1. J. Heikkila and O. Silven, in *Proceedings of IEEE Computer Vision and Pattern Recognition* 1106 (1997).

2. L. Stirling, K. Willcox, P. Ferguson, and D. Newman, Aviation, Space, and Environmental Medicine **80,** 522 (2009).
3. Y. Fujii and K. Shimada, Meas. Sci. Technol. **17,** 2705 (2006).
4. Y. Fujii, K. Shimada, M. Yokota, S. Hashimoto, Y. Sugita, and H. Ito, Rev. Sci. Instrum. **79,** 056105 (2008).
5. Y. Fujii and K. Shinmada, Trans. Jpn. Soc. Aeronaut. Space Sci. **50,** 251 (2008).
6. Y. Fujii, K. Shimada, K. Maru, M. Yokota, S. Hashimoto, N. Nagai, and Y. Sugita, Trans. Jpn. Soc. Aeronaut. Space Sci. Space Technol. Jpn. **7,** Th_1 (2009).
7. Y. Fujii, K. Shimada, and K. Maru, Microgravity Sci. Technol. **22,** 115 (2010).
8. R. Y. Tsai, IEEE J. Robotics and Automation **3,** 323 (1987).
9. J. Zhang, D. Zhu, and Z. Zhang, Acta Opt. Sin. (in Chinese) **28,** 1552 (2008).
10. D. Wu, N. Lu, and J. Ouyang, Acta Opt. Sin. (in Chinese) **28,** 482 (2008).
11. J. Wang, H. Hong, C. Wang, and Y. Zhu, Electronics Optic & Control (in Chinese) **14,** (4) 94 (2007).
12. R. Zhang and J. Liu, Chin. J. Electron Devices (in Chinese) **30,** 1618 (2007).
13. N. Yang, J. Wang, C. Huang, R. Wang, and D. Jin, Tsinghua Univ. (Sci. & Tech.) (in Chinese) **40,** (4) 24 (2000).
14. R. M. Haralick, C. Lee, K. Ottenberg, and M. Nolle, IJCV **13,** 592 (1994).
15. C. Su, Y.-Q. Xu, H. Li, S. Q. Liu, and D. G. Li, Chin. J. Computers (in Chinese) **21,** 1084 (1998).
16. R. Horaud, B. Conio, O. Leboulleux, and B. Lacollt, Computer Vision, Graphics, and Image Processing **47,** 33 (1989).
17. F.-C. Wu and Z.-Y. Hu, J. Software (in Chinese) **14,** 682 (2006).
18. D. Liang, F. Wu, Z. Ruan, and S. Wei, J. University of Science and Technology of China (in Chinese) **32,** 192 (2002).
19. M. A. Penna, Pattern Recognition **24,** 533 (1991).
20. T. Jin, H. Jia, W. Hou, Y. Ryo, N. Norihiro, F. Yusaku, M. Koichi, O. Naoya, and S. Kazuhito, "Study on extraction method of the circle outline and its center in a moving dummy mass" Electronics World (to be published).