IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 12, 2019, accepted July 27, 2019, date of publication August 2, 2019, date of current version August 15, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2932835

Pose Measurement for Non-Cooperative Target Based on Visual Information

LONGZHI ZHANG[1], DONGMEI WU[1], AND YUQI REN[1,2]
[1]State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150001, China
[2]Shenzhen Whalehouse Technology Company Ltd., Shenzhen 518000, China

Corresponding author: Dongmei Wu (wdm@hit.edu.cn)

This work was supported by the Self-Planned Task of the State Key Laboratory of Robotics and System under Grant SKLRS201712A.

ABSTRACT In the process of on-orbit servicing, chasers achieve approach and contact with target spacecraft are via a series of orbital transfer and motion control. However, most target spacecraft are non-cooperative targets. As a matter of fact, mainly technical difficulty of non-cooperative targets still lies in unable directly obtaining their motion parameters, thereby, it is impossible to realize their autonomously rendezvous and capture. Thus, measure motion parameters of non-cooperative targets from outside is significant for research on-orbit service. In this paper, a pose measurement method coupled initial pose measurement with model-based tracking is put forward for space non-cooperative targets. Specifically, an algorithm based on SIFT matching to measure initial pose of target tracking is proposed to start subsequent target tracking. Afterwards, a target tracking algorithm is presented, which is via *a priori* information and the combination of edge features and point features. As well as, an M estimator is introduced to improve the accuracy and robustness of target tracking algorithm. Finally, an experimental platform is constructed to fulfill experiments to validate the effectiveness of proposed method. Also, the experimental results demonstrate that our approach has highly precision and robustness in vary illuminations. Simultaneously, our method could satisfy the real-time requirement of pose measurement for non-cooperative targets.

INDEX TERMS Pose measurement, non-cooperative target, model-based tracking, feature fusion.

I. INTRODUCTION

Capturing non-cooperative targets is indispensable to multitudes of on-orbit missions, such as prolong spacecraft lifespan, assistance faulty spacecraft into pre-selected orbit, and push abandoned satellites into the grave orbit [1]–[3]. Thus, investigate on non-cooperative target capture has gained increasing attention from researchers. Yet due to state of motion is unknown, research difficulty on capturing non-cooperative targets still lies in accurate estimation of the motion parameters [4]–[7]. To address this issue, researchers come up with quite a few solutions.

DLR designed a capture mechanism for non-cooperative target orbit engines, and presented a complex non-cooperative target recognition algorithm based on multi-sensor data fusion [8]. However, this method was fully controlled from ground system. When the delay was large, entire system would be unstable, thereby the reliability of method would be affected. The United States carried out the FREND project mainly for non-cooperative spacecraft to perform capture, derailment and other operations, which used three hand-eye cameras as the pose measurement system [9]–[11]. In this project, real-time was critical to the success of missions. But coordination of multiple cameras and too involved processing loop constrained the real-time performance of tasks. Nassir Oumer and Panin [12] put forward a method for stereo camera-based 3D tracking to estimate pose of a non-cooperative satellite. In their approach, for each point, update stage of the filter was iterated until convergence, which leads to the pose estimation time would be affected by iterative times. Fourie *et al.* [13], [14] established a visual-inertial system based on stereo vision, and described a vision-based relative navigation and control strategy for inspecting an unknown non-cooperative object in space. Although this method has advantages of highly reliability and precision, time of visual information processing would be restricted by external environment. Segal *et al.* [15] developed a stereovision-based filtering algorithm for estimating the relative state between two non-cooperative spacecraft. They employed multiple iterated extend Kalman filters, and

The associate editor coordinating the review of this manuscript and approving it for publication was An-An Liu.

each endowed with a different hypothetical target inertia. The algorithm was robust to inertia modeling uncertainties, and could estimate the complete relative state, but estimation time was not fully considered. Peng *et al.* [16] constructed a stereo vision system to measure the relative pose of non-cooperative target. Compared with traditional methods, their scheme has highly accuracy and efficiency, but exists a problem of large calculation amounts. Ying *et al.* [17] put forward an effective method to identify the features of non-cooperative targets and track their pose based on point cloud. But the method has a large amount of computation, due to it was based on particle filter algorithm.

Although above solutions adopted different algorithms to achieve pose measurement, they have similarity in idea. Firstly, the employed computer vision method to imaging. Afterwards, image processing and pose algorithm were adopted to measure position and attitude parameters. Finally, the captured part were identified. Nevertheless, these schemes are defective, a quintessential example should be cited that large computational will give rise to system delay, thereby the real-time performance of method could not be guaranteed.

To deal with previously mentioned problem, a real-time pose measurement method for non-cooperative target is put forward, in condition of ensuring the precision and robustness. The remainder of this paper is organized as follows. Section II introduces the research motivation of our contributions. Then, section III gives an exhaustively formulation of our thought. Additionally, experimental results are shown in Section IV is to demonstrate the superiority of presented approach. Ultimately, in Section V, we conclude the paper, simultaneously look forward the future work.

II. RESEARCH MOTIVATION

Indeed, essence of pose measurement of three-dimensional object is to measure the translation and rotation parameters of world coordinate (model coordinate) fixed on object in camera coordinate. That is, pose measurement is to measure both target translations along three coordinate axis and rotation angles generated by target rotate around three coordinate axis. In mathematical, these translations and rotation angles are represented by homogeneous transformation matrix.

Yet non-cooperative targets unable to provide priori information of cooperation cursor and motion state used for detection and tracking, which brings certain difficulties to pose information acquisition. But, motion of non-cooperative targets in camera are continuous. Thus, priori information could be obtained from image frame in camera to reduce the calculations of pose measurement, which means that the pose information could be acquired in real-time through frame-by-frame sampling in video sequence. Since target motion is a continuous process, pose in previous frame image could provide multitudes of prior information for latter frame image. When utilizing these prior information in pose solution, not only pose measurement error will be reduced to a large extent, but also pose measurement time will be lessened [18]–[22]. Therefore, priori information of the first frame image during measuring non-cooperative targets pose is quite critical. Meanwhile, insufficient prior information indicates that measurement algorithm of non-cooperative target pose in initial frame image should be remarkable different from the subsequent algorithm.

Furthermore, this research is focus on the need for non-cooperative targets in near-earth space, and hence, real-time must be taken into consideration. Whereas, among computer vision methods used for non-cooperative target pose measurement, model-based method not only has good robustness in variable space environment, but also has lower computation and better real-time performance. It means that model-based method could deal with the rotation, irregular motion, etc., as well as could realize the real-time output of non-cooperative target pose information in the video sequence [23]–[31]. In addition, model-based method could reduce mismatch points to improve the pose measurement accuracy.

Last but not the least, compared with large devices, such as laser and radar, visual camera could provide stable and reliable 2D visual information, simultaneous has small space and low energy consumption. Accordingly, a visual camera is adopted as hardware device of pose measurement in this paper.

To sum up, a pose measurement method for non-cooperative target based on model is put forward. The proposed method is composed of two paramount components, respectively are initial pose measurement and model-based target tracking. Firstly, initial pose of subsequent target tracking is gained by performing pose measurement on a query image selected in video sequence. Then, initial pose is exploited to tracking non-cooperative target in video sequence with frame by frame to output the real-time pose. In order to attain the stable, real-time, and accurate output of non-cooperative target pose in final, pose tracking precision needs to be judged. When the accuracy conform to the demands, keep tracking continues, else, pose of non-cooperative target in current location needs to be recalculated, as a new initial pose to restart tracking.

III. METHOD

A. INITIAL POSE MEASUREMENT OF NON-COOPERATIVE TARGET TRACKING

Measure initial pose of target tracking aims at detecting and locating non-cooperative target in three-dimensional environment. To accomplish pose measurement in computer vision, our idea is via extract certain feature information of target, and matching these information with image information of target. Because of space target is sensitive to illumination, occlusion, and view-point variations, while SIFT feature extraction operator has better adaptability to these factors, an algorithm based on SIFT matching to measure initial pose of target tracking is proposed to initial the subsequent target tracking, which flow path is summarized as follows,

1. Measure pose in query image I_q as reference pose;

2. Extract SIFT feature points from query image I_q and current image I_c;

3. Eliminate interference points in I_q and I_c;

4. Perform keypoints matching between I_q and I_c by Brute-Force matcher;

5. Solve corresponding relationship between SIFT feature points in current image I_c and 3D points in non-cooperative target model based on reference pose;

6. Couple virtual visual servo with RANSAC (RANdom SAmple Consensus) to solve pose parameters in I_c and regard it as initial pose.

Briefly, in order to measure initial pose of target tracking, our algorithm has three important parts, respectively are, reference pose measurement of non-cooperative target in query image, extract and matching SIFT feature points between query image and current image, and pose solution based on multipoint matching.

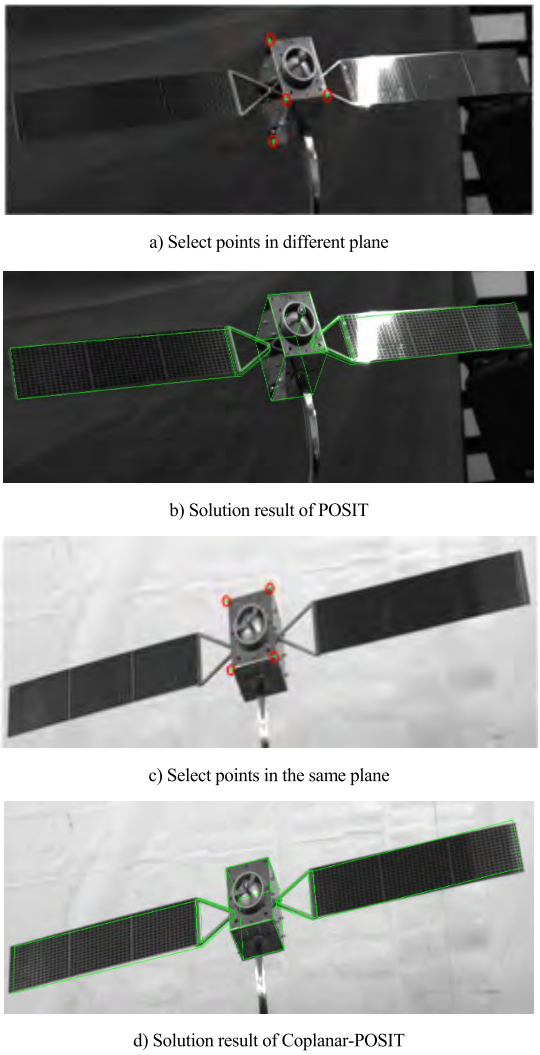
1) MEASURE REFERENCE POSE

In our implementation, pose parameters of non-cooperative target are known (with a certain uncertainty) at the first frame in a given video sequence. Thus, one frame is selected from video sequence as a query image, and pose parameters of non-cooperative target in this image is called reference pose. To obtain the reference pose, POSIT pose solving algorithm is employed owing to its lower computational complexity, faster convergence, and better robustness.

Nevertheless, classical POSIT could only perform pose calculation for the selected reference points in non-coplanar, for the selected reference points in coplanar, coplanar-POSIT is adopted to achieve pose solving [32]. Unlike POSIT, coplanar-POSIT could solve two sets of pose parameters. During iteratively solving pose parameters, above two sets of pose parameters are judged. Only the solution with less error is used for each iteration, which could make the final iteration converge results basically approximate in two sets. Thus, any of two data sets could be used as the final result of pose measurement.

For the sake of validating the reliability and accuracy in reference pose solution, we randomly extract a frame image as query image. Then, we utilize human-computer interaction to respectively input three-dimensional coordinates of four points and their corresponding two-dimensional image feature points, and determine whether they are coplanar. Finally, according to the judgment result, Posit or coplanar-Posit pose measurement algorithm is selected to solve position and attitude of non-cooperative target in query image. The solution result of reference pose is shown in figure 1.

As is exhibited in figure 1, we respectively select four non-coplanar points and four coplanar points from non-cooperative target image, then, non-cooperative target wireframe model is re-projected on image to evaluate accuracy of reference pose solution. Obviously, from figure 1, solution result of POSIT is better than coplanar-POSIT.



a) Select points in different plane

b) Solution result of POSIT

c) Select points in the same plane

d) Solution result of Coplanar-POSIT

FIGURE 1. **Solution results of reference pose.**

To sum up, in this work, image plane points corresponding to the feature points of known non-cooperative target model are selected through the means of human-computer interaction, as well as, after coplanar judgment of 3D feature points on model, POSIT or coplanar-POSIT is used to solve reference pose of non-cooperative target in query image.

2) EXTRACT AND MATCHING SIFT FEATURE POINTS

Since current image has certain translation, rotation, and scale transformation with respect to the query image, yet SIFT feature extraction operator has rotation invariance, scale invariance, as well as has certain resistance to changes in illumination. In this paper, SIFT feature extraction operator is adopted to extract SIFT feature points from query image and current image, furthermore, extracted feature points is used for matching, after eliminating feature points in background and mismatch, correspondence between 3D feature points on model and 2D SIFT feature points in current image is established by 2D-3D re-projection.

During pose measurement, extract SIFT feature points from query image is actually a process of learning query image. That is, the extracted feature points and their descriptors will be used as learning samples, to match SIFT feature points and descriptors in current image, and to establish a one-to-one mapping relationship between them.

Extracted SIFT feature points from non-cooperative target images is shown in figure 2.

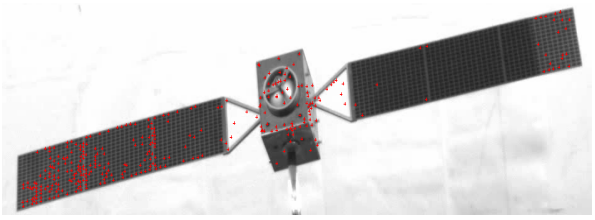


FIGURE 2. Extraction of SIFT feature points.

As previous mentioned, one-to-one mapping relationship between query image and current image needs to be built in SIFT feature points matching. Brute Force matching algorithm [33] is used for SIFT feature points matching because of its stable matching result and highly calculation efficiency. Specifically, after matching feature descriptor in query image with feature descriptor in current image to return the optimal match, and setting threshold to eliminate the mismatch, a one-to-one mapping relationship between query image and current image is established.

Matching result of SIFT feature points achieved by Brute Force as shown in figure 3.

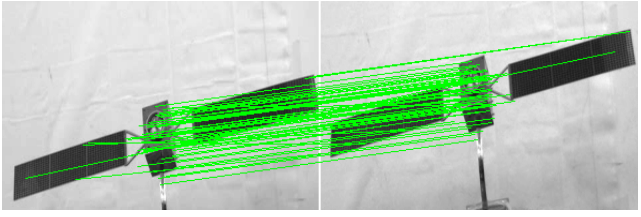
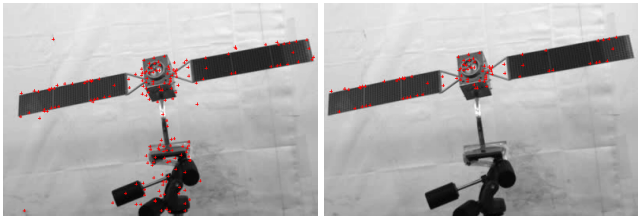


FIGURE 3. Matching result of SIFT feature points by brute force.

Moreover, during extracting SIFT feature points, some points located outside the non-cooperative target and belong to the background, named background interference points, are also extracted. Thus, the extracted SIFT feature points need to be screened.

Screening background interference points is on the basis of hypothesis that all SIFT feature points are located on the surface of the non-cooperative target model. Meanwhile, non-cooperative target model and reference pose are known, which could be utilized to eliminate SIFT feature points in background through projecting on 2D image. In fact, screening background interference points is to only retain SIFT feature points on non-cooperative target in image. The result of background interference points elimination as show in figure 4.



a) Extracted SIFT feature points b) Retained SIFT feature points

FIGURE 4. Result of eliminate background interference points.

After eliminating the background interference points, three-dimensional coordinates of points in model coordinate system corresponding to the remaining SIFT feature points need to be solved. These three-dimensional coordinates and their corresponding feature descriptors will be used for subsequent pose measurement based on multi-point matching.

Solution of three-dimensional coordinate is also based on the assumption that feature points are located on a certain plane of non-cooperative target model surface. Since non-cooperative target model and its reference pose in query image are known, in accordance with space projection principle, coordinates of two-dimensional image feature points in three-dimensional model coordinate could be easily obtained.

3) POSE MEASUREMENT BASED ON MULTIPOINT MATCHING

Along with obtaining the three-dimensional coordinates of feature points corresponding to the model coordinate system, correspondence between 3D coordinates of SIFT feature points on non-cooperative target and 2D feature points in image is also acquired, which could be used for solving non-cooperative target position and attitude of current image in camera coordinate system. This paper couples virtual vision servo [34] with RANSAC [35] to solve pose parameters in current image.

Basic idea of virtual vision servo for solving non-cooperative target pose is to regard the problem of solving pose parameters as the dual topic of 2D visual servo. In visual servo process, a particular target in image could be continuously observed by controlling the motion of camera, which could be achieved by minimizing the error between desired state parameter s^* of image feature and actual state parameter s. Under ideal conditions, exists a unique camera pose to minimize this error.

Indeed, large amounts of 3D feature points ${}^oP$ exist on measured target, that is, ${}^oP$ is 3D coordinates of points on measured target in model coordinate system. We set a virtual camera, and define its position and attitude in model coordinate system as r, then through minimize the error between observed data s^* (2D coordinate of SIFT feature points in image is p, at this time, $s^* = p$) and 2D coordinate s (according to pose parameter r, 3D feature points are forward projected to image plane to obtain s), named Δ, actual pose parameters of measured target relative to camera could

be acquired.

$$\Delta = \sum_{i=1}^{N} (s_i(r) - s_i^*)^2 = \sum_{i=1}^{N} (pr_\varepsilon(r, {}^oP_i) - s_i^*)^2 \quad (1)$$

where, $pr_\varepsilon(r, {}^oP_i)$ is projection model obtained based on camera internal parameter ξ and camera pose r, N refers amount of feature points.

During pose calculation, motion of virtual camera initially located at r_i position follows virtual visual servo control algorithm to minimize the error parameter Δ, at convergence position of algorithm, virtual camera reaches the r_d pose to minimize Δ, r_d is pose parameter to be solved.

In classical visual servo process, define task function as e,

$$e = (s(r) - s^*) \quad (2)$$

Take the derivative of formula 2,

$$\dot{e} = \frac{\partial e}{\partial s}\frac{\partial s}{\partial r}\frac{dr}{dt} = Lv \quad (3)$$

where, L represents image Jacques matrix, and v refers velocity vector.

Then, e is decoupled exponentially,

$$\dot{e} = -\lambda e \quad (4)$$

where, λ is proportional coefficient for controlling attenuation rate.

Thus, following control laws could be obtained,

$$v = -\lambda \hat{L}_s^+ e \quad (5)$$

where, $\hat{L}_s$ refers pseudo-inverse matrix of L, $\hat{L}_s^+ = (\hat{L}_s^T \hat{L}_s)^{-1} \hat{L}_s^T$.

Finally, velocity vector v is mapped to its corresponding instantaneous displacement, to update the pose parameters of non-cooperative target,

$${}^cM_o^{k+1} = {}^cM_o^k \exp(v) \quad (6)$$

where, $\exp(v)$ represents exponential mapping in the form of homogeneous matrices, k refers iteration numbers in an iterative optimization process.

Due to coarse error may be exist in the process of feature matching and 3D information solving, which will affect the accuracy of pose control algorithm based on virtual visual servo. In order to make full use of correspondence of 2D-3D SIFT feature points, eliminate possible coarse error, improve the accuracy and reliability of algorithm, RANSAC is introduced for pose measurement, namely, virtual visual servo is combined with RANSAC to solve non-cooperative target pose.

RANSAC assumes that data sample contains both correct data and abnormal data. It solves problem by repeatedly selecting a set of random subset in data sample, and employs acquired results to test other data to determine whether these data are interior points. At last, all interior points are used for resolving model to obtain the result as correct as possible.

TABLE 1. Algorithm of pose parameters solution.

Algorithm Couple Virtual Visual Servo with RANSAC to Solve Pose Parameters
Input feature points in non-cooperative target model and in their corresponding images
solve reference pose;
while inner points is insufficient ***and*** unattained the maximum iterations
randomly select 4 sets of data points;
POSIT algorithm is exploited to solve pose parameter r;
solve residual l based on pose parameters r;
if $l < threshold$
then
for each feature point ${}^OP_i$, ***do***
solve error between re-projection points of feature points in model based on pose parameter r and feature points in image;
$error = \left\| pr_\xi\left(r, {}^oP_i\right) - p_i \right\|$
if $error < threshold$
then
feature point is inner point;
end if
end for
end if
end while
in accordance with all inner points obtained via above steps and based on virtual visual stereo algorithm to resolve pose parameters of non-cooperative target;
Output pose of non-cooperative target in camera coordinate

In this paper, 2D-3D matching relationship of feature points is a data sample containing abnormal data. Firstly, four feature points are randomly selected to solve pose parameters. Afterwards, we make a decision on all data points in entire data point set, regard points which are close enough to projection points and feature points as interior points. Finally, pose of non-cooperative target is to be resolved in accordance with all obtained interior points. The algorithm of pose parameters solution via couple virtual visual servo with RANSAC as shown in table 1.

In above algorithm, reference pose is acquired by query image based on POSIT. However, if error between reference pose and solved pose is too large, the algorithm will diverge or converge at local minimum. At this time, it is essential to re-select an image close to current image as reference image, and solve its pose as reference pose, then, measure pose in current image as initial pose of following target tracking process, to start model-based target tracking algorithm.

B. NON-COOPERATIVE TARGET TRACKING

Since three-dimensional model of non-cooperative target is known, we employ model-based tracking to realize real-time

solution of non-cooperative target pose parameters. But in order to improve the accuracy of target tracking, a target tracking algorithm based on edge feature fusion point feature is proposed, which takes Harris corner features and edge features on non-cooperative target in images as tracking features. As well as, an M estimator is introduced to enhance the resistance of nonlinear optimization pose to external points. Finally, pose parameters could be real-time and accurate outputted.

1) SEARCHING AND MATCHING EDGE FEATURE

Edge is a prominent feature in image, which has advantages of stable extraction effect and strong anti-interference ability, but it is difficult to quantify. Thus, we transform edge features into sample point features with a certain distance on the edge, and then, perform a normal one-dimensional search to determine the matching between model edge points and image edge points. The details as follows,

1. According to pose parameter r_{k-1} of non-cooperative target in previous frame image I_{k-1}, 3D model of non-cooperative target is projected on image plane of current $I_k th$ frame image, to obtain two-dimensional projection edge of model;

2. Sampling for two-dimensional projection edges of model at a certain distance, to acquire a sampling point set of projection edge on model in the $I_k th$ frame image, which is defined as $\{x_i\}_{i=1}^{Ng}$;

3. For any sampling point x_i on projection edge of model, through the ECM searching algorithm [36] to perform one-dimensional search along the normal n_i of x_i point on model projection edge, and seeking the nearest gray gradient extreme point in image as corresponding point x_i', further to acquire the set of image feature points $\{x_i'\}_{i=1}^{Ng}$, which is correspond to the set of sampling points $\{x_i\}_{i=1}^{Ng}$.

The one- dimensional searching along normal as shown in figure 5.

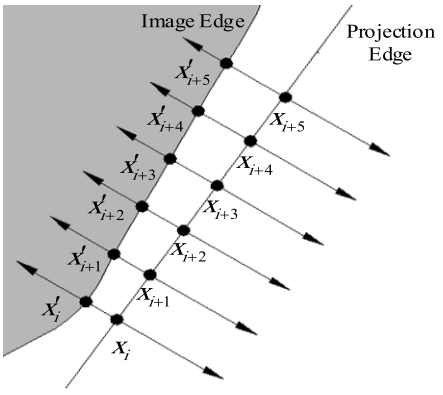


FIGURE 5. One-dimensional searching along normal.

In above normal one-dimensional searching process, an $m * m$ convolution template is to be set, which direction is determined in accordance with normal n_i. Moreover, within searching range R, convolution value of each possible pixel (x_{ij}) along normal n_i is obtained by the convolution template, which has a one-to-one correspondence to likelihood ratio ζ_j. Also, points with the maximum likelihood ratio in image and larger than the set threshold are recorded as corresponding points x_i', which will be used for following nonlinear optimization process of pose parameters.

It is well to be reminded that in process of tracking non-cooperative target, complex small contour features on non-cooperative target, as well as change in visibility of each surface during rotation, will greatly effect on the accuracy of tracking. Hence, to address this issue, this paper judges the visibility of non-cooperative target model plane, and sets the minimum threshold of projection line.

In fact, during the tracking process, only edge features on visible plane could perform low-order edge tracking to update the pose parameters. Whereas in this paper, determination of plane visibility is achieved by solving the included angle α, as shown in figure 6.

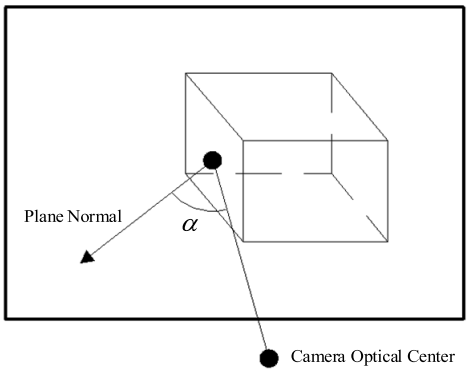


FIGURE 6. Determination of plane visibility.

As demonstrated in figure 6, α is an included angle between the line connecting camera optical center to gravity center of plane to be determined and plane normal. In this paper, the threshold k is defined as 80°, when $\alpha < 80°$, plane to be determined is visible, but when $\alpha > 80°$, plane to be determined is invisible.

On the other hand, in the first step of tracking process, non-cooperative target model needs to be projected on a two-dimensional plane for low-order edge tracking, yet in this process, line size determination of model projection is added to improve tracking accuracy. It is stipulated that the line is included in subsequent sampling matching process only when the model projection line length is greater than a certain threshold, and the small contour, which projection line length is less than this threshold is excluded from the tracking.

To sum up, edge feature matching points $\{x_i'\}_{i=1}^{Ng}$ in image as show in figure 7.

2) POINT FEATURE DETECTION AND TRACKING

Naturally, multiple similar edge features might be extracted, due to the complexity of projected image. In order to distinguish features, point feature is taken as another visual

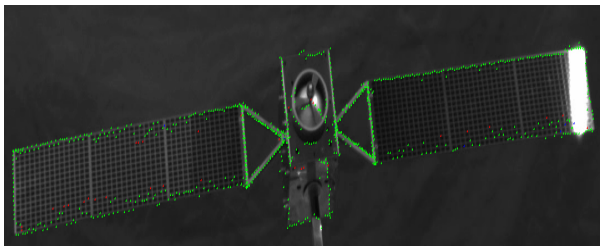


FIGURE 7. Edge feature matching points in image.

feature applied to tracking, since it also has strong anti-interference ability. In fact, target tracking using feature points is achieved by nonlinearly optimizing the objective function Δ^p (Δ^p is an objective function about the distance from feature point extracted in current frame to the feature point tracked in following frame) to solve current pose of the non-cooperative target. Consequently, this paper uses 3D feature points on non-cooperative target model and 2D feature points in image to achieve directly correspond 3D to 2D, and acquires pose parameter r via optimizing the objective function Δ^p.

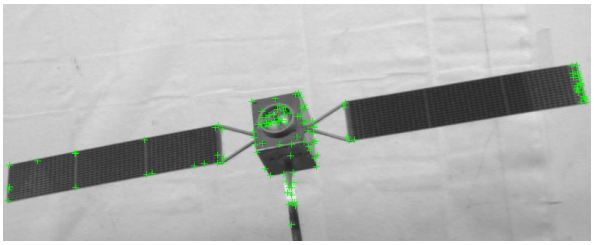


FIGURE 8. Harris corner feature of non-cooperative target.

In our research, point feature is detected by Harris corner detector [37] as shown in figure 8, and its tracking is realized via KLT tracking [38]. That is, our implement reduces the calculation of tracking by reducing number of tracking points, to improve the real-time performance of tracking accordingly.

In this paper, details in point tracking based on KLT as follows,

1. Extracting continuous single frame image from a video sequence of non-cooperative target;

2. Detecting non-cooperative target ontology from image sequence as foreground target;

3. Recording a certain frame from image sequence as I_O (foreground target appears in I_O);

4. Extracting feature points from I_O via Harris operator;

5. Selecting randomly two continuous frames as I_k and I_{k+1} from image sequence, their order are behind I_O (foreground target also appears in I_k and I_{k+1});

6. Searching position of feature points extracted from I_O in I_k, and solve their coordinate parameters;

7. Interacting for above 6 steps to complete KLT tracking of feature points.

In above process, define $\{x_i\}_{i=1}^{Np}$ as a feature point set extracted from the $I_k th$ frame image, then its corresponding 3D feature point set on non-cooperative target model is recorded as $\{X_i\}_{i=1}^{Np}$. Apparently, feature point set $\{x_i'\}_{i=1}^{N_p}$ in the $I_{k+1}th$ frame image correspond to $\{x_i\}_{i=1}^{Np}$ is obtained through solving optical flow from the $I_k th$ frame image to the $I_{k+1}th$ frame image, which has a one-to-one mapping relationship with $\{X_i\}_{i=1}^{Np}$.

Specially, during extracting Harris feature points, some background interference points are also the included in extracted feature point, which elimination ways is similar to the method mentioned above (remove background interference points from SIFT feature points extraction). Finally, only Harris feature points on non-cooperative target contours are retained.

3) TARGET TRACKING BASED ON FUSION OF EDGE FEATURE AND POINT FEATURE

As a matter of fact, accumulative error of pose would be generated during extracting and tracking of point features. Fusion of edge feature and point feature is to avoid the disadvantage of two features, and improve the accuracy and robustness of target tracking. In consequence, after acquiring correspondence between 2D feature points and 3D feature points, issue of solve pose parameter could be equivalent to a local nonlinear optimization problem. Through local nonlinear optimization of forward projection error Δr between above two features, the estimated value $\hat{r}$ of pose parameter r is finally obtained.

$$\hat{r} = \arg\min_{r} \Delta r \tag{7}$$

$$\Delta r = \sum (e_i(r))^2 \tag{8}$$

where, $e_i(r)$ is the error between two-dimensional feature and projection feature.

It can be seen from formula (8) that this is a nonlinear minimization problem about the pose parameter r. Therefore, Gauss-Newton iteration [39] is applied to iteratively update the pose parameter r, and finally, optimal estimation of r could be obtained.

Define set of projection points as $\{x_{i,j}\}$ acquired by projecting non-cooperative target model to image, set of points on non-cooperative target in image correspond to $\{x_{i,j}\}$ is $x'_{i,j}$, error function $\Delta(r)$ refers distance from points $x'_{i,j}$ in image to model projection points $x_{i,j}$, then,

$$\Delta(r) = \sum_i \sum_j (d_\perp(x_{i,j}, x'_{i,j}))^2 \tag{9}$$

where, $d_\perp(x_{i,j}, x'_{i,j})$ is distance from point $x'_{i,j}$ in image to model projection point $x_{i,j}$.

In our implement, starting from the initial pose estimation r_0, in each interaction loop k, performing a tiny displacement transformation δr for pose parameter r_k, and using following formula to update pose parameter,

$$r_{k+1} = r_k \oplus \delta r \tag{10}$$

where, $\oplus$ is sum operator in parameter space, which ultimate goal is to determine the value of r_k that minimizes the Δr in

each iteration loop, and its description in mathematics is,

$$\delta r = \arg\min_{\delta r} \sum_i \sum_j (d_\perp(x_{i,j}(r_{k+1}), x'_{i,j}))^2 = \arg\min_{\delta r} \sum_i \sum_j (d_\perp(x_{i,j}(r_k \oplus \delta r), x'_{i,j}))^2 \quad (11)$$

In process of interaction, tiny displacement transformation δr is acquired by following formula,

$$\delta r = -J^+ e(r) \quad (12)$$

where, J^+ is pseudo-inverse matrix of J, and J is Jacobian matrix of error vector e(r), $J = \partial e(r) / \partial r$.

Especially, tiny displacement transformation δr is screw displacement of pose parameter, and $\delta r = (v, w)$, where v refers translation displacement parameter, w represents rotation displacement parameter.

In order to achieve the update of pose parameters through tiny displacement transformation δr, converting the δr to its corresponding rigid body instantaneous displacement δM via an exponential map,

$$\delta M = \exp([\delta r]) \quad (13)$$

where, $[\delta r] = \begin{bmatrix} w & v \\ 0 & 0 \end{bmatrix}$.

In addition, define ${}^cM_o$ as homogeneous transformation matrix in process of rigid body transformation, which represents pose parameters of non-cooperative target relative to the camera.

Then, on the basis of formula (10), ${}^cM_o$ could be update by δM,

$$M_{k+1} = \exp([\delta r]) M_k \quad (14)$$

However, during solving pose parameters of each frame in video sequence by using Gaussian-Newton nonlinear optimization, the input parameters maybe contain some outliers, which will affect the accuracy of pose parameters obtained by iterative optimization. Consequently, M estimator [40] is introduced to eliminate the outliers may be generated during model projection and edge matching, due to its better robustness.

In this work, objective of M estimator is to reduce the influence of outliers in input parameters on estimation results, via assigning the adaptive weight to input parameter in error function. Therefore, define $_R$ as estimation function of pose error, then, the expression of nonlinear optimization issue as following,

$$\Delta_R = \rho(s(r) - s^*) \quad (15)$$

where, $\rho(e_i(r))$ is robust estimation function.

In above formula, diagonal weight matrix D ($D = diag(w_1, \ldots, w_n)$) is determined by setting the M estimator ρ, as well as, D is used for setting the weight w_i of each parameter in error vector e under the condition of given error vector $e = [e_0\ e_1\ \ldots\ e_n]$, where $0 < w_i < 1$, and w_i is expressed as,

$$\omega_i = \frac{\psi(\delta_i/\sigma)}{\delta_i/\sigma} \quad (16)$$

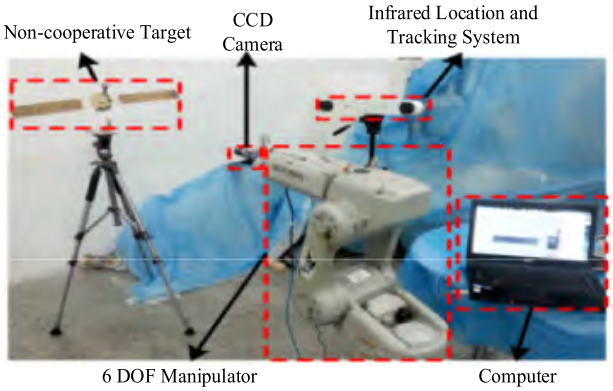


FIGURE 9. Experimental platform.

where, δ_i is normalized residuals ($\delta_i = e_i - Med(e)$), Ψ is influence function, Med refers median operator, and σ represents standard deviation of interior point data.

Especially, the influence function Ψ is defined as,

$$\psi(u) = \begin{cases} u(C^2 - u^2)^2| & u| \le C \\ 0| & u| > C \end{cases} \quad (17)$$

where, C is constant.

Also, standard deviation of interior point data σ is another important parameter, which is unknown in initial case, but it could be obtained by median absolute deviation $\hat{\sigma}$,

$$\hat{\sigma} = \frac{1}{\Phi^{-1}(0.75)} Med_i(|\delta_i - Med_j(\delta_j)|) \quad (18)$$

where, $\Phi()$ is normal cumulative distribution function, and $\Phi^{-1}(0.75) = 1.48$.

As a result, after above calculation, the formula (12) could be updated,

$$\delta_r = -(DJ)^+ De(r) \quad (19)$$

Additionally, in process of solving non-linear pose parameters mentioned above, the error function is also need to be solved and non-linear optimization.

Firstly, error function of edge feature is to be solved, which is expressed as,

$$\Delta^g(\boldsymbol{r}) = \frac{1}{N_g} \sum_i \rho^g(e_i^g(\boldsymbol{r})) = \frac{1}{N_g} \sum_i \rho^g(\sigma_g^{-1} d_\perp(l_i(\boldsymbol{r}), \boldsymbol{x}'_i)) \quad (20)$$

where, $d_\perp(l_i(r), x'_i)$ represents distance from feature point x'_i to projection straight line $l_i(r)$, ρ^g is M estimator, and σ^g refers normalization factor of standard deviation for e_i^g, which is expressed as $\sigma_g = \sqrt{\frac{1}{N_g} \sum_i \rho^g(d_\perp(l_i(r_f), x'_i))}$.

Obviously, function expression of $l_i(r)$ needs to be known to obtain the $d_\perp(l_i(r), x'_i)$.

Define function expression of initial model projection straight line $l_i(r_k)$ as,

$$x \cos\theta_{x_i} + y \sin\theta_{x_i} = \rho_{x_i} \quad \forall (x, y) \in l_i(r_k) \quad (21)$$

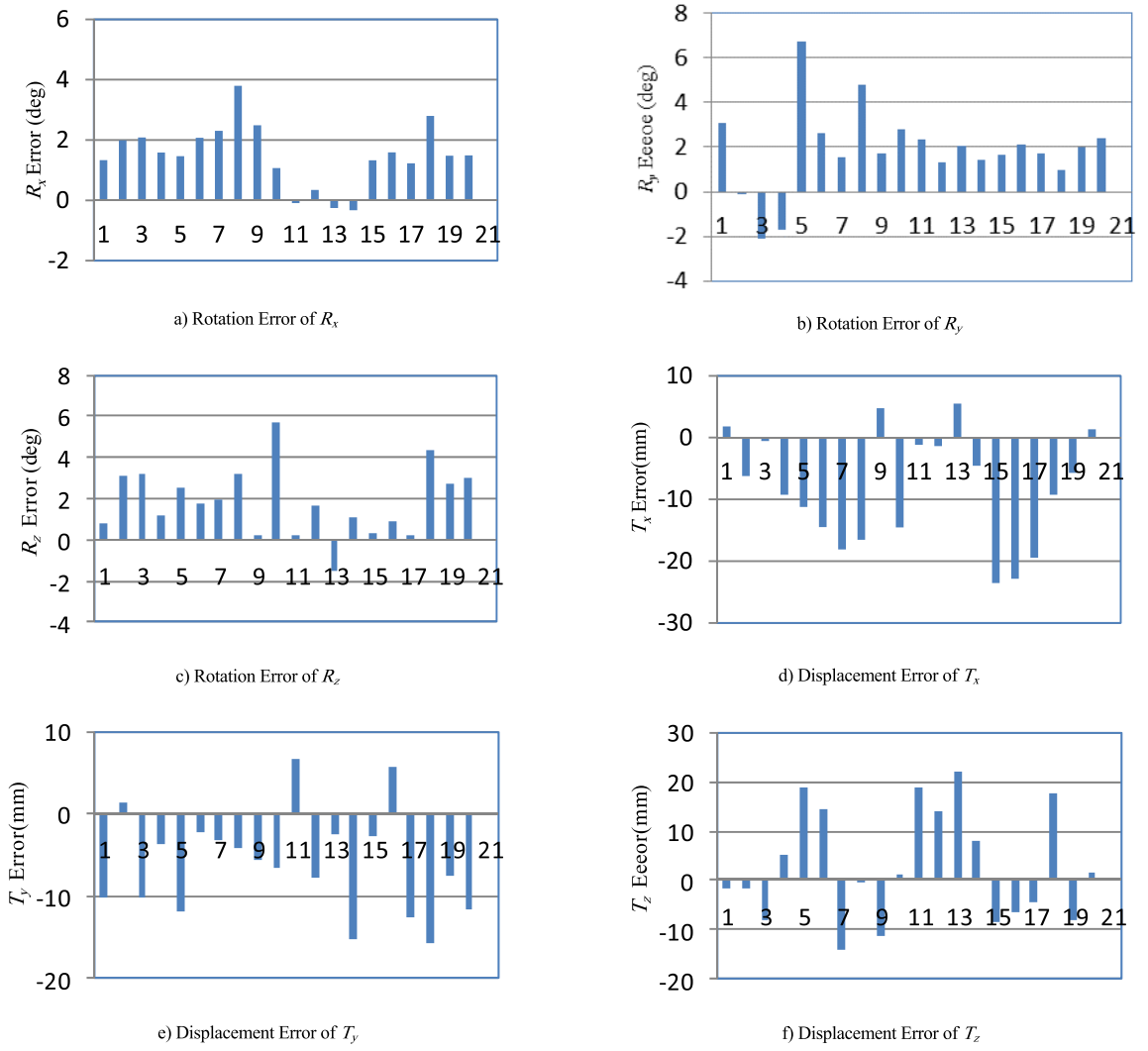


FIGURE 10. Distribution of absolute pose measurement error.

where, θ_{x_i} is included angle between x axis of image plane coordinate and normal direction of $l_i(r_k)$, ρ_{x_i} refers distance between $l_i(r_k)$ and image plane coordinate origin.

Since θ_{x_i} and ρ_{x_i} could be acquired in process of model projection, as well as pose parameters in previous frame of non-cooperative target model are known, then, function representation of straight ling L_i in model coordinate on three-dimensional model of non-cooperative target corresponding to $l_i(r_k)$ could be acquired.

In accordance with current pose parameter r, $l_i(r)$ could be obtained by projecting L_i to image plane and update θ_{x_i} and ρ_{x_i}. Then, $d_\perp(l_i(r), x_i')$ could described as,

$$d_\perp(l_i(r), x_i') = \rho_{x_i} - \rho_{x_i'} \tag{22}$$

In formula (22),

$$\rho_{\boldsymbol{x}_i'} = x_{\boldsymbol{x}_i'} \cos\theta_{\boldsymbol{x}_i} + y_{\boldsymbol{x}_i'} \sin\theta_{\boldsymbol{x}_i} \tag{23}$$

where, $x_{x_i'}$ and $y_{x_i'}$ are coordinates of x_i' in image plane coordinate.

Furthermore, Jacobian matrix $J_{e_i^g}$ of $e_i^g(r)$ needs to be solved, which is defined as,

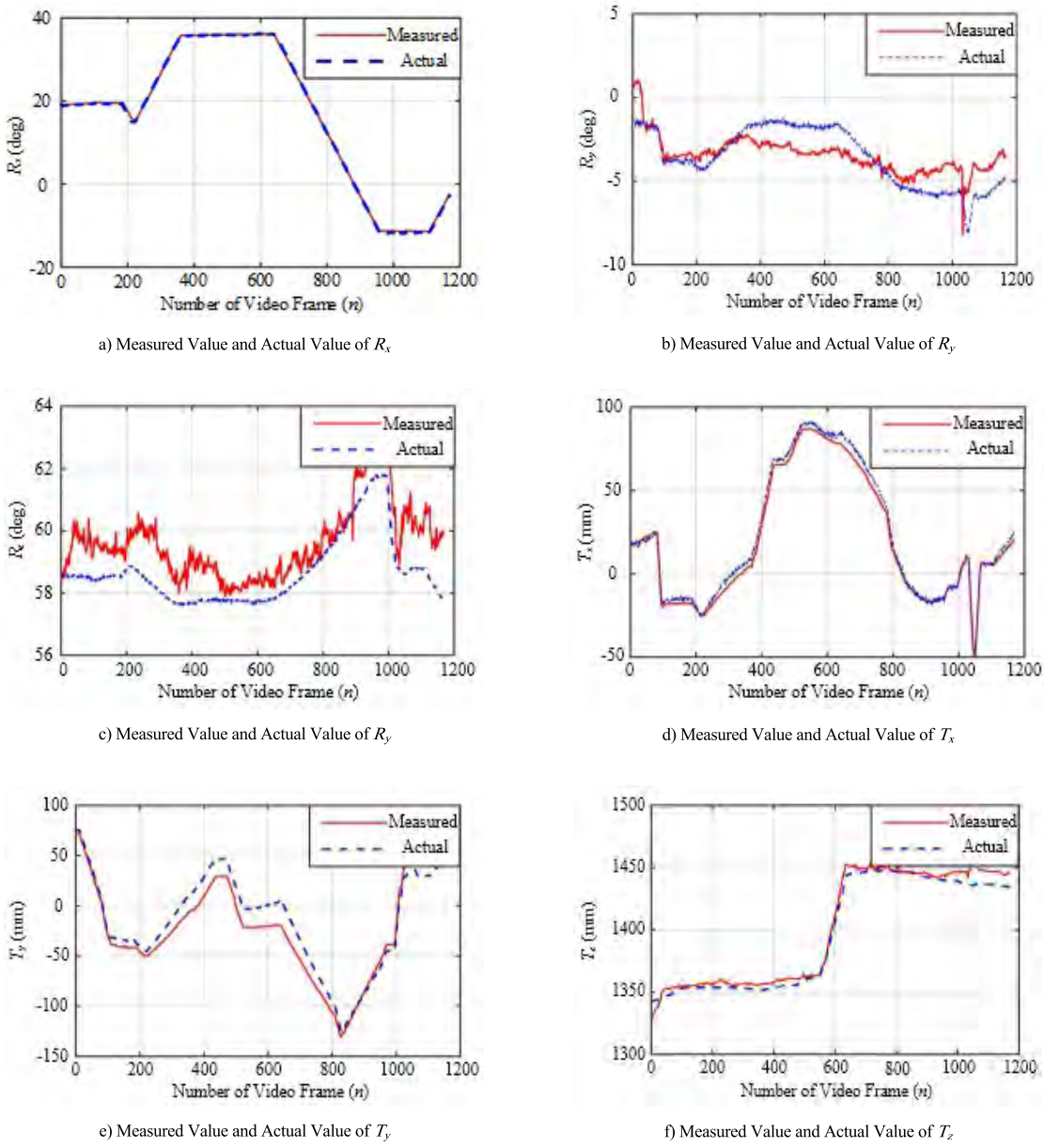
$$J_{e_i^g} = \frac{\partial e_i^g}{\partial r} = \frac{\partial e_i^g d_\perp(l_i(r), x_i')}{\partial r} \tag{24}$$

Then, substituting formula (17) to formula (24),

$$\begin{aligned} J_{e_i^g} &= \frac{\partial \rho_{\boldsymbol{X}_i}}{\partial r} - \frac{\partial \rho_{\boldsymbol{X}_i'}}{\partial r} \\ &= \frac{\partial \rho_{\boldsymbol{X}_i}}{\partial r} + (x_{\boldsymbol{X}_i'} \sin\theta_{\boldsymbol{X}_i'} - y_{\boldsymbol{X}_i'} \cos\theta_{\boldsymbol{X}_i'}) \frac{\partial \theta_{\boldsymbol{X}_i}}{\partial r} \\ &= \frac{\partial \rho_{\boldsymbol{X}_i}}{\partial r} + \alpha \frac{\partial \theta_{\boldsymbol{X}_i}}{\partial r} \\ &= J_{\rho_{x_i}} + \alpha J_{\theta_{x_i}} \end{aligned} \tag{25}$$

where, $\alpha = x_{X_i'} \sin\theta_{X_i'} - y_{X_i'} \cos\theta_{X_i'}$, $J_{\rho_{x_i}}$ and $J_{\theta_{x_i}}$ is obtained by literature [41],

$$\begin{aligned} J_{\rho_{x_i}} &= [\lambda_{\rho_{x_i}} \cos\theta_{x_i} \lambda_{\rho_{x_i}} \sin\theta_{x_i} - \lambda_{\rho_{x_i}} \rho_{x_i} \cdots \\ &\quad \times (1 + \rho_{x_i}^2) \sin\theta_{x_i} - (1 + \rho_{x_i}^2) \cos\theta_{x_i} 0] \\ J_{\theta_{x_i}} &= [\lambda_{\theta_{x_i}} \cos\theta_{x_i} \lambda_{\theta_{x_i}} \sin\theta_{x_i} - \lambda_{\theta_{x_i}} \rho_{x_i} \cdots \end{aligned} \tag{26}$$



a) Measured Value and Actual Value of R_x

b) Measured Value and Actual Value of R_y

c) Measured Value and Actual Value of R_y

d) Measured Value and Actual Value of T_x

e) Measured Value and Actual Value of T_y

f) Measured Value and Actual Value of T_z

FIGURE 11. Pose parameter cure of target tracking.

$$\times \rho_{x_i} \cos\theta_{x_i} - \rho_{x_i} \sin\theta_{x_i} - 1] \tag{27}$$

In formula (27),

$$\lambda_{\rho_{x_i}} = \frac{(A\rho_{x_i} \cos\theta_{x_i} + B \sin\theta_{x_i} + C)}{D} \tag{28}$$

$$\lambda_{\theta_{x_i}} = \frac{(A\rho_{x_i} \sin\theta_{x_i} + B \cos\theta_{x_i} + C)}{D} \tag{29}$$

In formula (28) and formula (29), $AX+BY+CZ+D=0$ is a function representation of a certain three-dimensional plane to which the straight line L_i belongs in model coordinate. Since model parameters of non-cooperative target are known, four parameters (A, B, C and D) of three-dimensional plane could be obtained by using pose parameters of previous frame image and three-dimensional coordinates of mentioned sampling point X_i in model coordinate.

Through combining formulas (26), (27), (28) and (29), the needed Jacobian matrix $J_{e_i^g}$ for nonlinear optimization of pose parameters could be obtained,

$$J_{e_i^g} = \begin{bmatrix} \lambda_{d_\perp} \cos\theta_{x_i} \\ \lambda_{d_\perp} \sin\theta_{x_i} \\ -\lambda_{d_\perp} \rho_{x_i} \\ (1+\rho_{x_i}^2) \sin\theta_{x_i} - \alpha\rho_{x_i} \cos\theta_{x_i} \\ -(1+\rho_{x_i}^2) \cos\theta_{x_i} - \alpha\rho_{x_i} \sin\theta_{x_i} \\ -\alpha \end{bmatrix}^T \tag{30}$$

where, $\lambda_{d_\perp} = \lambda_{\rho_{x_i}} + \alpha\lambda_{\theta_{x_i}}$.

Secondly, according to one-to-one mapping relationship between three-dimensional point set $\{X_i\}_{i=1}^{N_p}$ and two-dimensional point set $\{x_i'\}_{i=1}^{N_p}$, error function of point feature

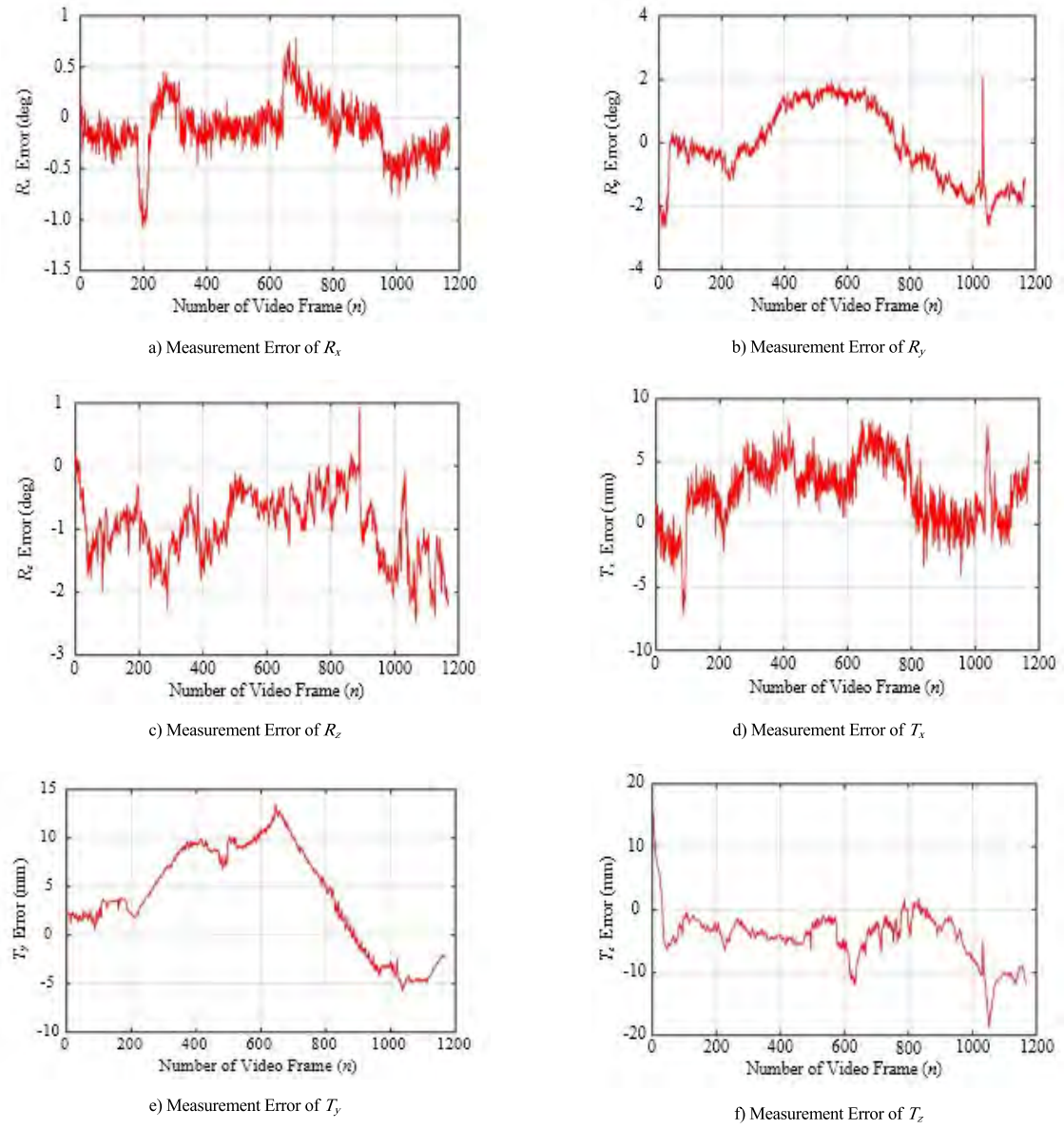


FIGURE 12. Error curve of target tracking.

could be expressed as follows,

$$\Delta^p(r) = \frac{1}{N_p}\sum_{i}^{N_p}\rho^p(e_i^p) \tag{31}$$

where, ρ^p is M estimator, and e_i^p is defined as,

$$e_i^p = \sigma_p^{-1}(x_i(r) - x_i') \tag{32}$$

where, $x_i(r)$ represents the result of three-dimensional point X_i being projected into image plane based on pose parameter r to be solved, namely, $x_i(r) = pr(X_i, r)$, σ_p is normalization factor of standard deviation for e_i^p, which is expressed as $\sigma_p = \sqrt{\sum_i \rho^p(x_i(r_f) - x_i')\Big/ N_p}$.

Moreover, Jacobian matrix Je_i^p of point feature is also acquired by literature [41],

$$Je_i^p = \frac{\partial e_i^p}{\partial r} = \begin{bmatrix} p_x & 0 \\ 0 & p_y \end{bmatrix} \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -1-x^2 & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & 1+y^2 & xy & -x \end{bmatrix} \tag{33}$$

where, p_x and p_y are focal ratio parameters of camera, x and y are coordinates of feature point $x_i(r) = pr(X_i, r)$ in image, Z refers depth value of 3D point on model corresponding to feature point x_i.

Lastly, in order to achieve target tracking based on fusion of edge feature and point feature, thereby to solve pose parameters with nonlinear optimization algorithm, a global error function Δ composed of edge feature error function and point feature error function is set up,

$$\Delta = \omega^g \Delta^g + \omega^p \Delta^p \tag{34}$$

where, Δ^g is error function of edge feature, Δ^p is error function of point feature, w^g and w^p are weight parameter, meanwhile, $0 < w^g < 1, 0 < w^p < 1$.

Similarly, global error vector e is described as,

$$e = \begin{bmatrix} \sqrt{\lambda^g} e^g \\ \sqrt{\lambda^p} e^p \end{bmatrix} \tag{35}$$

where, $e^g = \begin{bmatrix} e_1^g \\ \vdots \\ e_{N_g}^g \end{bmatrix}, e^p = \begin{bmatrix} e_1^p \\ \vdots \\ e_{N_p}^p \end{bmatrix}$.

As well as, global Jacobian matrix is expressed as,

$$J = \begin{bmatrix} \sqrt{\lambda^g} J_{e^g} \\ \sqrt{\lambda^p} J_{e^p} \end{bmatrix} \tag{36}$$

where, $J_{e^g} = \begin{bmatrix} J_{e_1^g} \\ \vdots \\ J_{e_{Ng}^g} \end{bmatrix}, J_{e^p} = \begin{bmatrix} J_{e_1^p} \\ \vdots \\ J_{e_{Np}^p} \end{bmatrix}$.

In formula (35) and formula (36), λ^g and λ^p are described as,

$$\lambda^g = \frac{\omega^g}{N_g}, \quad \lambda^p = \frac{\omega^p}{N_p} \tag{37}$$

However, it is known that from formula (19), global weight matrix D also needs to be solved for performing Gauss-Newton optimization based on M estimator, which is expressed as,

$$D = blockdiag(D^g, D^p) \tag{38}$$

Accordingly, the formula (19) could be updated as,

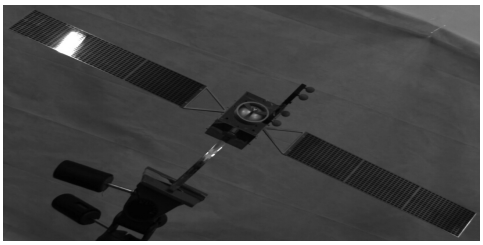
$$\begin{aligned} \delta_r &= -(DJ_s)^+ De(r) \\ &= -(\lambda^g J^{gT} D^{gT} D^g J^g + \lambda^p J^{pT} D^{pT} D^p J^p)^{-1} * \\ &\quad \times (\lambda^g J^{gT} D^{gT} D^g e^g(r) + \lambda^p J^{pT} D^{pT} D^p e^p(r)) \end{aligned} \tag{39}$$

Finally, through using above formula and Gauss-Newton iterative optimization algorithm, pose parameters of non-cooperative target could be solved and update in real-time. As a consequence, target tracking based on fusion of edge feature and point feature is exploited for experimental performance assessment.

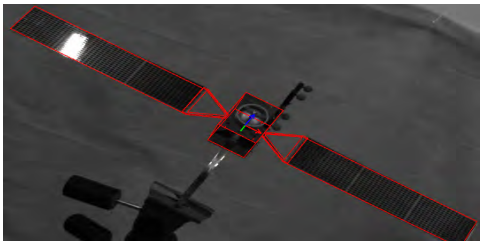
IV. EXPERIMENTS

A. ESTABLISH EXPERIMENTAL PLATFORM

Objective of experiment is to verify the effectiveness of proposed method in this paper. Thus, in accordance with research needs, we make use of existing conditions and equipment to establish our experimental platform, as shown in figure 9.



a) Imaging Effect of Poor Imaging



b) Target Tracking Effect of Poor Imaging

FIGURE 13. Target tracking experiment under poor imaging condition.

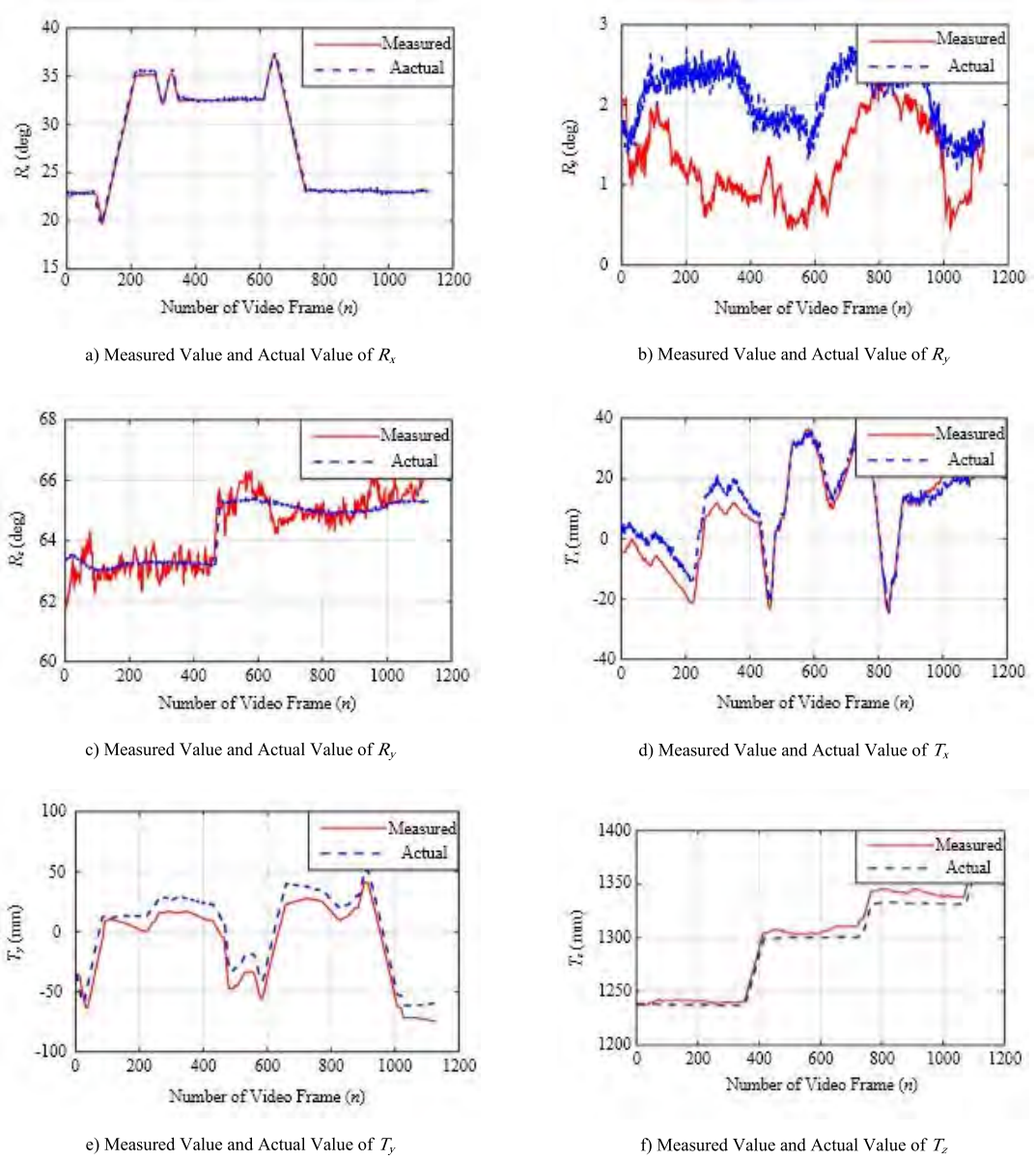
It can be clearly seen from figure 9 that established experimental system is composed of non-cooperative target model, CCD camera, 6 DOF manipulator, infrared location and tracking system, and a computer. More precisely, adopted non-cooperative target model is Chang'e-2 satellite model with a ratio of 1:30 compared to the real satellite. The CCD camera transmits collected image data to computer is via network port communication. Meanwhile, type of 6 DOF manipulator is a motorman SV3X industrial manipulator with repeated positioning accuracy of 0.03mm and a maximum working radius of 677mm. In our experiments, CCD camera is fixed to the end of manipulator through a special fixture. It means that motion of CCD camera follows movement of manipulator. Thus, manipulator is used for simulating motion of non-cooperative target relative to CCD camera. Furthermore, infrared location and tracking system is Polaris Spectra System produced by NDI Company of Canada, which transmits actual pose parameters through the USB interface. The transmitted data is regarded as reference data to evaluate the accuracy of proposed algorithm in this paper. Last but certainly not least, computer is used for collecting image data and solving pose parameters.

Additionally, it is worth noting that in our experiments, since measurement range of Polaris Spectra System and view field range of CCD camera are limited, position range of pose measurement experiment is limited from 1200mm to 1550mm, which is equivalent to measure non-cooperative target pose at a distance from 36m to 46.5m via using ratio of 1:30 for conversion.

B. EXPERIMENTAL RESULTS AND ANALYSIS

1) MEASUREMENT ACCURACY VERIFICATION OF INITIAL POSE

In order to evaluate the accuracy of proposed initial pose measurement algorithm, we collect 20 groups of



a) Measured Value and Actual Value of R_x

b) Measured Value and Actual Value of R_y

c) Measured Value and Actual Value of R_y

d) Measured Value and Actual Value of T_x

e) Measured Value and Actual Value of T_y

f) Measured Value and Actual Value of T_z

FIGURE 14. Pose parameters curve of tracking in condition of poor imaging.

experimental data, and analyze error between measurement values and actual values. Specially, measurement values are obtained by the proposed algorithm, and actual values are acquired via Polaris Spectra System. For convenience, we define error between measurement value and actual value as absolute pose measurement error, T_x, T_y, and T_z represent translations along x, y, and z axes, respectively, meanwhile, R_x, R_y, and R_z respective represent rotations along x, y, and z axes.

The distribution of absolute pose measurement error is shown in figure 10.

In this experiment, we use the average value of absolute pose measurement error to evaluate the measurement accuracy, as shown in table 2.

TABLE 2. Comparison of absolute pose measurement error.

Name	R_x	R_y	R_z	T_x	T_y	T_z
Average Error (deg/mm)	1.557	2.256	1.989	9.610	7.359	10.593

It is apparent from table 2 that average error of rotation parameters are no more than 2.5°, and average error of translation parameters are less than 10.6mm, which illustrates that in our implement, besides the systematic error of Polaris Spectra System, error of proposed initial pose measurement algorithm is relative to slight.

In our actual application, rotation precision is required to be less than 5°, and translation precision is required to be

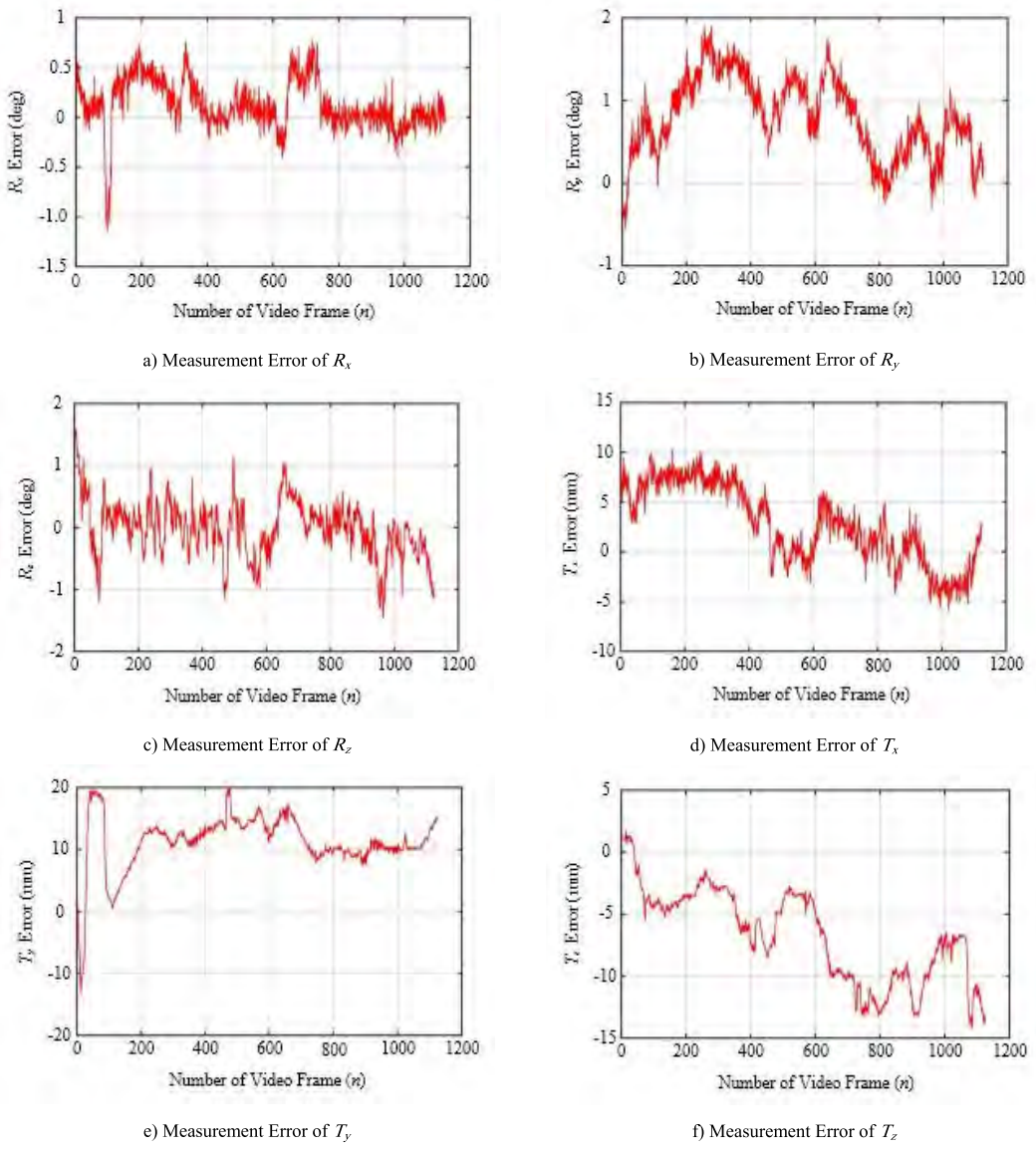


FIGURE 15. Pose error curve of tracking.

less than 20mm. Thus, the table 2 demonstrates that accuracy of initial pose measurement is highly enough to satisfy the requirement in actual application.

2) TARGET TRACKING

In order to verify the accuracy, robustness and real-time of model-based target tracking algorithm, as well as further to examine the reliability and accuracy of initial pose obtained by pose measurement algorithm after starting the target tracking, we designed and conducted model-based target tracking experiments.

Primarily, target tracking experiment in condition of normal light is performed, to validate feasibility and accuracy of proposed pose measurement algorithm. Comparison images between actual values acquired by experiment and measured values are shown in figure 11, as previous mentioned, T_x, T_y, and T_z represent translations along x, y, and z axes, respectively, meanwhile, R_x, R_y, and R_z respective represent rotations along x, y, and z axes.

As is exhibited in figure 11 that red curve is measured value, and blue curve represents the actual value. Afterwards, error curve of target tracking could be acquired via data processing for figure 11 is shown in figure 12.

As is known from analysis of figure 11 and figure 12, error of rotation tracking along x, y, and z axes are less than 2.5°, and error of translation tracking along x, y, and z axes are less than 20mm, which satisfy the requirement of precision in our actual application. Thus, the experimental results demonstrate that proposed tracking algorithm has an ideal tracking effect.

Next in line, target tracking experiment under poor imaging condition is accomplished, to confirm robustness and accuracy of proposed pose measurement algorithm. The imaging effect and target tracking effect as shown in figure 13, pose parameter curve of tracking is shown in figure 14, and pose error curve of tracking is displayed in figure 15. Also, T_x, T_y, and T_z represent translations along x, y, and z axes, respectively, meanwhile, R_x, R_y, and R_z respective represent rotations along x, y, and z axes.

It can be seen form figure 14 and figure 15 that rotation tracking accuracy of rotation parameters R_x, R_y, and R_z are less than 2°, and tracking accuracy of translation parameters T_x, T_y, and T_z are less than 20mm. Obviously, accuracy of tracking under poor imaging condition is not significantly different from normal condition, which illustrate that proposed pose tracking algorithm has good robustness, and could satisfy the pose tracking requirements under different lighting conditions to some extent.

Finally, tracking speed verification experiment of pose tracking algorithm is carried out, to test the pose parameters calculation speed of tracking algorithm. In this implement, the algorithm is used for tracking poses of four video sequences with different frames, and the recorded data is shown in table 3.

TABLE 3. Data of pose tracking speed verification experiment.

Video Sequence Number	Frames Number in Video (frame)	Calculation Time (s)	Processing Speed (frame/s)
1	1439	139.5	10.32
2	1394	132.7	10.50
3	1125	114.1	9.86
4	1168	117.7	9.92

Table 3 reflects that pose tracking algorithm in this paper has a processing speed of about 10frame /s. In this work, image processing speed is required to be at least 4frame /s. Thus, table 3 illustrate that our algorithm could satisfy the requirements of real-time pose tracking for non-cooperative targets in real-time.

V. CONCLUSION AND DISCUSSIONS

To solve the pose measurement problem of non-cooperative targets in space environment, a pose measurement approach combined initial pose measurement with model-based target tracking is put forward. Purpose of measure initial pose is to start target tracking. Thus, an algorithm based on SIFT matching to measure initial pose of target tracking is proposed, which is to initial the target tracking. To solve the pose parameters, the algorithm couples virtual visual servo with RANSAC based on multipoint matching. Then, a target tracking algorithm is proposed. Distinct from conventional model-based tracking, edge and point are taken as visual features in this paper. Moreover, an M estimator is introduced to resist the interference from environment during target tracking, to realize the real-time measurement of pose parameters. Lastly, a series of experiments are carried out to confirm the pose measurement method in this paper has highly precision, robustness, and good real-time performance. The experimental results demonstrate that our proposed method not only could adapt various illumination conditions, but also could satisfy the requirements of real-time solving pose parameters in video sequences.

Particularly, a beneficial work to be finished in the future is study on recover the structure of non-cooperative target from motion via using visual sensors, to realize the identification and pose measurement of unknown non-cooperative targets. Also, with the development of visual sensors, initial pose measurement will be further deeply investigated, such as based on RGB-D sensor, multi-lens, laser radar, and etc. In addition, how to predict and correct the data in pose tracking to improve the accuracy and robustness of tracking algorithm is an important topic of future research.

ACKNOWLEDGMENT

The authors would like to express the sincere gratitude to anonymous reviewers and editors for their professional suggestions on improving the manuscript. They also would like to thank Dr. Baoshi Cao for paying his precious time to put forward the valuable suggestion on writing.

REFERENCES

[1] N. G. Cui, P. Wang, J. F. Guo, and X. Cheng, "A review of on-orbit servicing," *J. Astronaut.*, vol. 28, no. 4, pp. 805–811, 2007.

[2] B. Liang, D. Xiaodong, L. Cheng, and X. Wenfu, "Advances in space robot on-orbit servicing for non-cooperative spacecraft," *Robot*, vol. 34, no. 2, pp. 242–256, Mar. 2012.

[3] A. Flores-Abad, O. Ma, K. Pham, and S. Ulrich, "A review of space robotics technologies for on-orbit servicing," *Prog. Aerosp. Sci.*, vol. 68, pp. 1–26, Jul. 2014.

[4] K. Landzettel, C. Preusche, A. Albu-Schaffer, and D. Reintsema, "Robotic on-orbit servicing—DLR's experience and perspective," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2006, pp. 4587–4594.

[5] F. Sellmaier, T. Boge, J. Spurmann, S. Gully, T. Rupp, and F. Huber, "On-orbit servicing missions: Challenges and solutions for spacecraft operations," in *Proc. Spaceops Conf.*, Apr. 2010, p. 213.

[6] M. H. Shan, J. Guo, and E. Gill, "Review and comparison of active space debris capturing and removal methods," *Prog. Aerosp. Sci.*, vol. 80, pp. 18–32, Jan. 2016.

[7] R. Tomasz, "Obstacle avoidance in space robotics: Review of major challenges and proposed solutions," *Prog. Aerosp. Sci.*, vol. 101, pp. 31–48, Aug. 2018.

[8] G. Hirzinger, K. Landzettel, B. Brunner, M. Fischer, C. Preusche, D. Reintsema, A. Albu-Schäffer, G. Schreiber, B.-M. Steinmetz, "DLR's robotics technologies for on-orbit servicing," *Adv. Robot.*, vol. 18, no. 2, pp. 139–174, 2004.

[9] T. Debus and S. Dougherty, "Overview and performance of the front-end robotics enabling near-term demonstration (FREND) robotic arm," in *Proc. AIAA Infotech*, 2009, pp. 1–12.

[10] J. Lennon, C. Henshaw, and W. Purdy, "An architecture for autonomous control of a robotic satellite grappling mission," in *Proc. AIAA Guid., Navigat. Control Conf. Exhibit*, 2013, p. 7259.

[11] J. Obermark, G. Creamer, B. E. Kelm, W. Wagner, and C. G. Henshaw, "SUMO/FREND : Vision system for autonomous satellite grapple," *Proc. SPIE*, vol. 6555, May 2007, Art. no. 65550Y.

[12] N. W. Oumer and G. Panin, "3D point tracking and pose estimation of a space object using stereo images," in *Proc. Int. Conf. Pattern Recognit.*, Nov. 2012, pp. 796–800.

[13] D. Fourie, B. Tweddle, S. Ulrich, and A. Saenz-Otero, "Vision-based relative navigation and control for autonomous spacecraft inspection of an unknown object," in *Proc. AIAA Guid., Navigat., Control (GNC) Conf.*, 2013, p. 4759.

[14] D. Fourie, B. E. Tweddle, S. Ulrich, and A. Saenz-Otero, "Flight results of vision-based navigation for autonomous spacecraft inspection of unknown objects," *J. Spacecraft Rockets*, vol. 51, no. 6, pp. 2016–2026, May 2014.
[15] S. Segal, A. Carmi, and P. Gurfil, "Stereovision-based estimation of relative dynamics between noncooperative satellites: Theory and experiments," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 2, pp. 568–584, Mar. 2014.
[16] J. Peng, W. Xu, and H. Yuan, "An efficient pose measurement method of a space non-cooperative target based on stereo vision," *IEEE Access*, vol. 5, pp. 22344–22362, 2017.
[17] Y. He, B. Liang, J. He, and S. Li, "Non-cooperative spacecraft pose tracking based on point cloud feature," *Acta Astronautica*, vol. 139, pp. 213–221, Oct. 2017.
[18] J. Thienel, J. Van Eepoel, and R. Sanner, "Accurate state estimation and tracking of a non-cooperative target vehicle," in *Proc. AIAA Guid., Navigat., Control Conf. Exhibit*, 2006, p. 6802.
[19] L. Regoli, K. Ravandoor, M. Schmidt, and K. Schilling, "On-line robust pose estimation for rendezvous and docking in space using photonic mixer devices," *Acta Astronautica*, vol. 96, pp. 159–165, Mar./Apr. 2014.
[20] T. Tzschichholz, T. Boge, and K. Schilling, "Relative pose estimation of satellites using PMD-/CCD-sensor data fusion," *Acta Astronautica*, vol. 109, pp. 25–33, Apr./Jun. 2015.
[21] S. Sharma and S. D'Amico, "Comparative assessment of techniques for initial pose estimation using monocular vision," *Acta Astronautica*, vol. 123, pp. 435–445, Jun./Jul. 2016.
[22] V. Pesce, R. Opromolla, S. Sarno, M. Lavagna, and M. Grassi, "Autonomous relative navigation around uncooperative spacecraft based on a single camera," *Aerosp. Sci. Technol.*, vol. 84, pp. 1070–1080, Jan. 2019.
[23] R. Opromolla, G. Fasano, G. Rufino, and M. Grassi, "A review of cooperative and uncooperative spacecraft pose determination techniques for close-proximity operations," *Prog. Aerosp. Sci.*, vol. 93, pp. 53–72, Aug. 2017.
[24] F. Terui, "Model based visual relative motion estimation and control of a spacecraft utilizing computer graphics," in *Proc. 21st Int. Symp. Space Flight Dyn.*, Tolouse, France, Sep. 2009, pp. 1–15.
[25] B. Tamadazte, E. Marchand, S. Dembélé, and N. Le Fort-Piat, "CAD model-based tracking and 3D visual-based control for MEMS microassembly," *Int. J. Robot. Res.*, vol. 29, no. 11, pp. 1416–1434, Jul. 2010.
[26] V. Kumar, H. B. Hablani, and R. Pandiyan, "Relative navigation of satellites in formation using monocular model-based vision," *IFAC Proc. Volumes*, vol. 47, no. 1, pp. 497–504, 2014.
[27] R. Opromolla, G. Fasano, G. Rufino, and M. Grassi, "Performance evaluation of 3D model-based techniques for autonomous pose initialization and tracking," in *Proc. AIAA Infotech. Aerosp.*, 2015, p. 1426.
[28] R. Opromolla, G. Fasano, G. Rufino, and M. Grassi, "A model-based 3D template matching technique for pose acquisition of an uncooperative space object," *Sensors*, vol. 15, no. 3, pp. 6360–6382, Mar. 2015.
[29] L. Liu, G. Zhao, and Y. Bo, "Point cloud based relative pose estimation of a satellite in close range," *Sensors*, vol. 16, no. 6, p. 824, Jun. 2016.
[30] R. Opromolla, G. Fasano, G. Rufino, and M. Grassi, "Pose estimation for spacecraft relative navigation using model-based algorithms," *IEEE Trans. Aerosp. Electron. Syst*, vol. 53, no. 1, pp. 431–447, Feb. 2017.
[31] S. Sharma, J. Ventura, and S. D'Amico, "Robust model-based monocular pose initialization for noncooperative spacecraft rendezvous," *J. Spacecraft Rockets*, vol. 55, no. 6, pp. 1414–1429, Jun. 2018.
[32] G. Schweighofer and A. Pinz, "Iterative pose estimation using coplanar feature points," *Comput. Vis. Image Understand.*, vol. 63, no. 3, pp. 495–511, 1996.
[33] J. Cheng, C. Leng, J. Wu, H. Cui, and H. Lu, "Fast and accurate image matching with cascade hashing for 3D reconstruction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1–8.
[34] A. I. Comport, E. Marchand, M. Pressigout, and F. Chaumette, "Real-time markerless tracking for augmented reality: The virtual visual servoing framework," *IEEE Trans. Vis. Comput. Graph.*, vol. 12, no. 4, pp. 615–628, Jul./Aug. 2006.
[35] M. A. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[36] P. Bouthemy, "A maximum likelihood framework for determining moving edges," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 5, pp. 499–511, May 1989.
[37] C. G. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vis. Conf.*, Aug. 1988, vol. 15, no. 50, pp. 147–152.
[38] J. Shi and C. Tomasi, "Good features to track," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* Ithaca, NY, USA: Cornell Univ., Jun. 1993, pp. 593–600.
[39] S. Gratton, A. S. Lawless, and N. K. Nichols, "Approximate Gauss–Newton methods for nonlinear least squares problems," *SIAM J. Optim.*, vol. 18, no. 1, pp. 106–132, 2007.
[40] P. J. Huber, "Robust statistics," *J. Amer. Stat. Assoc.*, vol. 78, no. 381, pp. 1248–1251, 2011.
[41] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Autom.*, vol. 8, no. 3, pp. 313–326, Jun. 1992.

LONGZHI ZHANG received the B.S. degree in mechanical engineering from the Heilongjiang University of Science and Technology, in 2010, and the M.S. degree from Harbin Engineering University, China, in 2014. She is currently pursuing the Ph.D. degree with the State Key Laboratory of Robotics and System, Harbin Institute of Technology (HIT), China. Her research interests include computer vision, image processing, and robotics.

DONGMEI WU received the B.S. and M.S. degrees in computer science and automation engineering from Heilongjiang University, China, in 1990 and 1993, respectively, and the Ph.D. degree in mechatronics engineering from the Harbin Institute of Technology (HIT), in 2003, where she is currently a Professor with the State Key Laboratory of Robotics and System. Her research interests include computer vision, image processing, SLAM, and robotics.

YUQI REN received the B.S. and M.S. degrees in mechatronics engineering from the Harbin Institute of Technology (HIT), China, in 2013 and 2015, respectively. He is currently an Assistance Researcher with Shenzhen Whalehouse Technology Company Ltd., China. His research interests include computer vision, SLAM, and robotics.