Visual Motion Estimation for Tumbling Satellite Capture

M. Zampato, R. Pistellato, D. Maddalena, I. Bruno
Tecnomare S.p.A. S.Marco 3584, Venezia Italy
G. Visentin
ESA-ESTEC Noordwijk - The Netherlands

Abstract

Robotic service vehicles are presently seen as a viable means for performing satellite maintenance and ESA (European Space Agency) is presently investigating the possible characteristics and potential of a Geostationary Service Vehicle (GSV). A stereovision tracking system should be included in this GSV to help the operator track and capture the target satellite. This paper deals with a system, called TV-Trackmeter, previously developed for underwater robotics, featuring contemporary tracking and 3D measurements of more points. It is capable of providing rigid body real time pose evaluation and it is fitted with algorithms to recover the lost tracked targets. Trials have been carried out in the ESA-ESTEC premises, in order to evaluate the applicability of this concept to the satellite capture. A satellite mock-up has been moved by an industrial manipulator to simulate real motion of a free object in space. The results have proved the soundness of the approach, featuring for the linear position accuracy in the order of ±3 [mm] @ 3.5 [m] for the x and z axes and ±5 [mm] @ 3.5 [m] for the y axis. As regards the angular position error about the y axis is the smaller one, about ±0.5 [deg], while it is ±1 [deg] on the remaining axes.

1. Introduction

The commercial utilisation of the geostationary orbit (GEO) has nowadays a great importance from an economic and strategic point of view. Huge capital investments are currently made to develop, launch and manage this kind of satellites, which must guarantee suitable profits in their fields of applications, e.g. telecommunications, TV broadcasting, weather forecasting, etc. Therefore the possibility of performing maintenance of these satellites in order to solve problems and to increase the operational life is strongly required.

In principle this possibility has been demonstrated many times by direct astronaut intervention in Extra Vehicular Activity (EVA), for instance recently with the Hubble Space Telescope, which was built with the intention of being serviced later. On the other hand physical, technical and economic constraints of such an intervention make servicing by astronauts, on the existing fleet of conventional satellites, very unlike or even impossible.

Experiments were conceived in the past years aimed at evaluating the actual robotic capabilities in space interventions. One of the most important of them is ROTEX (see [1] for the basic concepts). In the definition of that experiment the grasping of a

BMVC 1996 doi:10.5244/C.10.68

free-flying object was included, even though there were no real-time video measuring devices yet available at that time. The experiment took place in 1993 and the flighting object grasping was one of the most stunning trials carried out [2] and was based on the approach described in [3].
Therefore robotic service vehicles are presently seen as viable means to face the satellite maintenance needs, and ESA is presently investigating the possible characteristics and potential of a Geostationary Service Vehicle (GSV) [4].
Basic tasks for such a vehicle consist of:

- chasing of the target satellite up to 100 m distance
- flying around with global inspection closing at about 50 m distance
- final approach up to ca. 3 m where the docking/berthing device is activated to capture the satellite.

Computer vision processing could help the operator carry out the control tasks, particularly as regards the final approach when the target satellite is an uncontrolled tumbling object. In this case the capture task can become difficult for the remote operator and an automatic means for zeroing the relative motion between GSV and target would be highly beneficial, in order to decrease time, costs and risks of this operation. For this purpose the computer vision system has to provide the GSV navigation control with an estimation of the rate of relative position and attitude between GSV and target, being the task for the control to get this motion equal to zero. Furthermore it is worthwhile to remark that existing satellites are not fitted with on purpose co-operative targets, such as backreflectors. What's worst most of their surface is normally covered with thermal blankets, entailing dazzling and very quickly variable images and thus providing very poor pictures for processing.
This paper deals with an experiment to test the efficiency and reliability of an improved real-time approach to estimate the pose of a flying object without any cooperative target. It is based on a stereovision system able to measure and track up to 6 patches simultaneously so as to asses the motion rate of a tumbling satellite. Such a system is the evolution of an existing system (the "TV-Trackmeter"), developed for underwater telerobotic applications, and already tested for space applications by ASI (the Italian Space Agency). The paper is arranged as follows: the next section describes the system, in terms of principle of operation, error analysis and system performances. Section 3 reports the satellite motion modelling, while section 4 deals with the results obtained during the trials. Finally section 5 draws some conclusions and presents the basic guidelines for further developments.

2. System Description

The "TV-Trackmeter" vision system includes:

- a stereo rig fitted with a pair of TV cameras
- three VME boards, a graphic monitor and a TV monitor. The boards include a Sparc10-MP CPU and two Frame Grabbers Eltec IPP.

The main features of the system are reported in Tab. 1.

Fig. 1. The stereo rig

TV Cameras	Sony CCD B&W fitted with auto iris lens
Resolution	560 (H) x 575 (V) TV lines
Focal length (Field of view)	12 mm (35° x 25°) Possibility of using as well: 4.8 mm (60° x 45°) 8 mm (45° x 30°)
Distance between cameras	280 [mm]
Frame grabbers resolution	726 (H) x 512 (V) Pixels, 8 bit/pixel

Tab. 1. Main Features of the stereo system

The basic functions of the "TV-Trackmeter" are described in the following sections.

A. 3D measurements

The operator selects the point to be measured by moving a cursor overlaid on the left camera image and the 3D euclidean coordinates of the point are calculated following a neighbourhood based stereo matching with a 8x8 pixel patch. The measure error standard deviation depends on the focal length used and in the worst case is in the order of 4 [mm] @ 2.5 [m]. The time required to get the measure is about 15 [ms].
Such accuracy is got through subpixel matching and a complex calibration routine based on a multiple-step linear/non-linear Levenberg-Marquardt optimisation.

B. Tracking

The system can also track the measured point at a tracking/measurement rate of 12.5 [Hz]. The tracking is neighbourhood based, so that the system does not need to use co-operative targets. To get better tracking performances a Kalman filter is implemented to cope with tracking speed as high as 200 [pixels/s]. It is also possible to track up to 6 points at the same time, at a lower rate of 6.25 [Hz], which allows for the motion estimation of rigid body (at least 3 point are to be tracked). The "TV-Trackmeter" is provided also with two options for the automatic selection of the best features to track based on either a Wavelet-Gabor directional filtering or the centroid point.

C. Motion Estimation
The motion estimation includes the position and the attitude of the tracked rigid body. The attitude information may be given in many ways (i.e. through a matrix of the SO(3)[1] group, α,β,γ angles or a quaternion [6]). "TV-Trackmeter" evaluates the latter because such a representation is singularity free.
We can distinguish 3 cases:
a) no geometric model is available
b) an apriori model is available
c) a geometric model can be defined by the "TV-Trackmeter" on the basis of a suitable set of measures

In case a) a reference frame is built as soon as the euclidean coordinates of 3 points referred to the stereo camera reference frame are available, as follows: the origin is fixed on the first point, the second point defines the X axis, and the third one defines the X-Y plane. The centroid of these points and the quaternion bound to the SO(3) map related to the reference frame so built yields the motion description. No geometric information of the object is needed, but there is no information for the tracking recovery once a target is lost. Nevertheless the system can still provide pose information, thanks to the redundant (more than 3) tracked point set, by locking onto another target.
In case b) the system is to be provided with the coordinates of at least 3 points in the object reference frame and the larger is the polygon formed by the tracked targets, the more accurate is the motion estimation. The tracking recovery is always possible onto the lost targets by applying the map from the object reference frame onto the TV cameras frame, being this map obtained through the current quaternion and centroid.
In case c) no geometric information is available, but the system can define a possible object model, in the TV cameras reference frame, by using the target measurements during a time interval in which the object is kept still. The output map is thus referred to the frozen reference frame the system has acquired at the start of the tracking phase.

D. Tracking recovery
In certain conditions (e.g. bad lighting conditions, temporary target occlusion) tracking of the single target point may be lost. The system can then try to retrieve the lost targets depending on the available geometric knowledge of the object.
When no geometric information is available, the system policy is to keep looking for a good new target in the same position where the old one was lost.
Provided that the geometric information is available or acquired at the beginning of the job, the system has got all the information to keep estimating the object motion when at least 3 targets are tracked. On the basis of that, the system evaluates the updated retinal coordinates of the lost points and tries to lock onto them again.

E. Data Filter and accuracy
The accuracy of the motion estimation is strictly bound to the precision of the "TV-Trackmeter" single measurement (see foregoing sections). Therefore, a low pass IIR filter has been integrated within the system so as to smooth the outputs by rejecting the higher frequencies due to the measure jitter and keeping the lower frequencies that own the motion information.

F. The error analysis

[1] SO(3) is the Special Orientation group of order 3. See [5].

The basic stereoscopy equations are:

$$\begin{cases} x = \dfrac{u}{f} iz \\ y = \dfrac{v}{f} jz \\ z = \dfrac{wfd}{s} \end{cases} \qquad (1)$$

where: (x,y,z) is the euclidean measure of a 3D point in the space, (i,j) is its retinal projection onto the reference image, f is the focal length, s is the displacement between the horizontal retinal projection of the stereo pair, d is the baseline between the cameras and u, v and w are system constants.

From (1) you can obtain the error propagation due to the uncertainty on the subpixel displacement between the stereo pair:

$$\begin{cases} \Delta x = -\left(\dfrac{u}{dwf^2}\right) iz^2 \Delta s \\ \Delta y = -\left(\dfrac{v}{dwf^2}\right) jz^2 \Delta s \\ \Delta z = -\left(\dfrac{1}{dwf}\right) z^2 \Delta s \end{cases} \qquad (2)$$

The error Δx and Δy depends on the retinal position of the measured target. It is then possible for them to fix an upper bound, being these errors larger near the screen border and null at the screen centre (i=0, j=0). As Δx and Δy have got relationships alike to Δz, being only a scale factor the difference, the probability distribution of Δz can account for all of them.

G. The stochastic random error

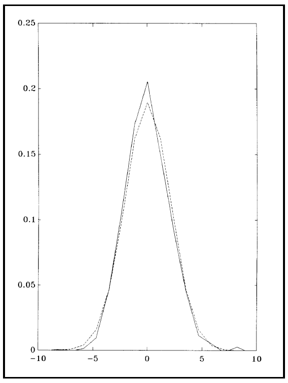
The first step is to evaluate the series expansion of the moment generatricx function ξ of the stochastic variable **z**, because you have from [7] the following relationship:

$$f_{\mathbf{z}}(a) = \mathcal{F}[\xi_{\mathbf{z}}(z)] \qquad (3)$$

where $f_{\mathbf{z}}(a)$ is the probability density of the stochastic variable **v**, and $\mathcal{F}$ is the Fourier transform functional. Given the ergodicity of the stochastic process **z**(t), the first 200 central moments of **z** have been gathered and by using a suitable series develop $\xi_{\mathbf{z}}(z)$ has been found. Finally, (3) yields $f_{\mathbf{z}}(a)$.

In Fig. 2 the theoretical result is shown by the solid line along with the dashed line that is the best Gaussian fitting.

Once the subpixel accuracy is fixed (as the hardware jitter and the algorithms are fixed), all the errors in (2) are $O(z^2)$ and the same holds for the standard deviation. This behaviour is summarised in Fig. 3 where the actual standard deviation is reported by a solid line, while the dashed line is the best fitting parabola. From the real standard deviation values and putting the system constants in (2), we have obtained the actual subpixel accuracy of the system that is currently 0.07 pixel, in practice not dependent on the range.



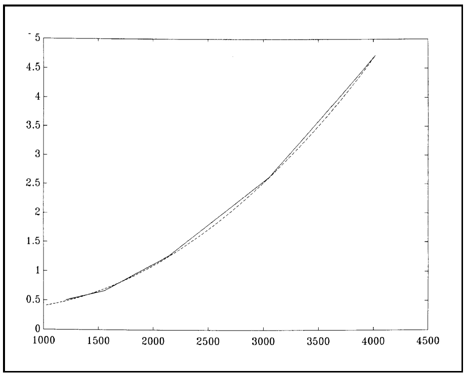


Fig. 2. Probability density error

Fig. 3. Parabolic behaviour of σ *versus the range.*

H. Tracking speed
The system is able to track high speed target thanks to a modified kinematic Extended Kalman Filter (EKF) [8] [9]. This kind of filter is widely described in literature [10]. The improvement the system gets by using such a filter concerns the wider tracking speed range. The EKF filter provided with a modified non-linear acceleration (2th order) model allows for not limited tracking speed even though it is limited in practice by the image blurring. A reliable upper bound of the tracking speed has turned out to be in the order of 200 [pixels/s].

3. Satellite modelling

The satellite modelling is the key point so as to determine is the vision system is suitable to track it and which accuracy is obtained.
Following what introduced in the previous sections, the satellite is moving as a free rigid body. The first step is then to get the dynamics motion equations. That is:

$$\mathbf{N} = \dot{\mathbf{L}} + \omega \times \mathbf{L} \tag{10}$$

where: $\mathbf{N}$ is net applied torque, $\mathbf{L}$ is the total angular momentum and ω is the angular speed of the rigid body in the solidal reference frame.
Furthermore, the basic relationship between $\mathbf{L}$ and ω is:

$$\mathbf{L} = \mathrm{I}\omega \tag{11}$$

$$\text{where: } \mathbf{I} = \begin{bmatrix} \mathrm{I}_{xx} & \mathrm{I}_{xy} & \mathrm{I}_{xz} \\ \mathrm{I}_{yx} & \mathrm{I}_{yy} & \mathrm{I}_{yz} \\ \mathrm{I}_{zx} & \mathrm{I}_{zy} & \mathrm{I}_{zz} \end{bmatrix} \qquad \textit{and}: \qquad \mathrm{I}_{xy} = \mathrm{I}_{yx} \quad \mathrm{I}_{xz} = \mathrm{I}_{zx} \quad \mathrm{I}_{yz} = \mathrm{I}_{zy} \tag{12}$$

If the body is free, there is no net torque. Moreover, fixing the body frame to the principal inertial axes, the inertial tensor **I** becomes diagonal and the dynamics equation yields:

$$\begin{cases} \dot{\omega}_x = \dfrac{I_{yy} - I_{zz}}{I_{xx}} \omega_y \omega_z \\ \dot{\omega}_y = \dfrac{I_{zz} - I_{xx}}{I_{yy}} \omega_z \omega_x \\ \dot{\omega}_z = \dfrac{I_{xx} - I_{yy}}{I_{zz}} \omega_x \omega_y \end{cases} \tag{13}$$

which are the so called Euler equation of the free rigid body motion. These equations are a differential system, whose motion integral brings about $\omega(t)$. It is worthwhile noting that if the body is symmetric, that is equivalent to say that two principal inertial momenta are equal, the motion of the rigid body is a precession motion around symmetry axis with variable nutation angle.

Once $\omega(t)$ is retrieved, the second step is getting a representation of the body attitude. Using quaternion it is possible to avoid kinematic singularities and get a straightforward differential relationship whose motion integral is the satellite attitude.

Let:

$q = \begin{bmatrix} q_s \\ q_x \\ q_y \\ q_z \end{bmatrix}$ the quaternion representation and $\dot{q} = \begin{bmatrix} \dot{q}_s \\ \dot{q}_x \\ \dot{q}_y \\ \dot{q}_z \end{bmatrix}$ its derivative

$\omega = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}$ the angular speed in the solidal reference frame and $V(\cdot) \rightarrow V(q) = \begin{bmatrix} q_x \\ q_y \\ q_z \end{bmatrix}$ an operator to get the vector part of the quaternion. The theory developed by Hamilton [6] about the screws and the finite displacements of the rigid bodies gives:

$$\omega = 2V(q^{-1}\dot{q}) = 2 \cdot \begin{bmatrix} -q_x & q_s & q_z & -q_y \\ -q_y & -q_z & q_s & q_x \\ -q_z & q_y & -q_x & q_s \end{bmatrix} \cdot \begin{bmatrix} \dot{q}_s \\ \dot{q}_x \\ \dot{q}_y \\ \dot{q}_z \end{bmatrix} \tag{14}$$

This relationship along with the constraint $|q|^2 = 1$ also yield:

$$\dot{q} = \frac{1}{2} \begin{bmatrix} -q_x & -q_y & -q_z \\ q_s & -q_z & q_y \\ q_z & q_s & -q_x \\ -q_y & q_x & q_s \end{bmatrix} \cdot \omega = \frac{1}{2} \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix} \cdot q \tag{15}$$

This is again a differential system that yields the quaternion q all along the motion.

4. Trials

The theory given here above has been tested in practice at the ESA-ESTEC premises by using: a COMAU manipulator to simulate the satellite motion, a satellite clung to the end effector and the "TV-Trackmeter" vision system to detect the satellite motion through machine vision.
The first step has been to evaluate the manipulator set points on the basis of the equations given in chapter 3 so as to model the free rigid body motion. From those equations the satellite motion is about a precession motion, as the body is symmetric in practice. This allows the vision system to track body neighbourhoods that never disappear during the capture phase, even though the shapes of the satellite features are rather mutable, owing to both the precession and nutation of the body. It is worth noting that the basic non-co-operative approach allows to cope with the problem with no apriori model available and with neighbourhoods that might even run out of visibility. Nevertheless, this sort of motion also backs the use of the model based approach, if any, that is even more robust, as the modelled and tracked features are always in sight all along the path.
The experiment aim was to determine the accuracy of the such a motion evaluation through the "TV-Trackmeter" system, in terms of position and speed, both translational and angular. The imposed motion was then checked with that measured by the device.
The measure error of the translational quantities (position and speed) is straightforwardly got through the difference between the imposed and the measured motions. The same holds for ω that can be evaluated through the "TV-Trackmeter" measurements by using (14). A little bit more tricky is instead the error evaluation for the attitude information as it is mapped with no singularities only by either a SO(3) matrix or a quaternion. In the latter case, the whole attitude error is simply given by:

$$\theta_{tot} = 2 \arcsin \sqrt{V(q_{err}) \cdot V(q_{err})^*} \qquad \textit{where}: \quad q_{err} = q_{vision} \cdot q_{arm}^* \tag{16}$$

This amount is the angular twist referred to the SO(3) map expressed by an equivalent axis of rotation and a twist about that. This does not allow to determine if the attitude error mostly happens about one specific cartesian axis. However, the 3 components of the vector:

$$\theta_{tot} \cdot \frac{V(q_{err})}{\left|V(q_{err})\right|} \tag{17}$$

allow for that. Note that these components cannot be thought at all of as angular rotations to be applied sequentially as a kinematic chain.
The evaluation of the errors was all carried out after an accurate inter calibration between the arm and the TV cameras reference frames. The result of such a step is a SE(3) matrix that allows for map the imposed and the measured motions onto the same reference frame fixed, for instance, on the end effector of the arm.
The tests have been carried out under several light conditions, taking into account that in the space environment they are quite harsh owing to the shadings, the reflections and the occlusions. The vision system was displaced with respect to the manipulator base reference frame along the y axis and aiming at that direction to the satellite.
A sample of the results is given hereafter.

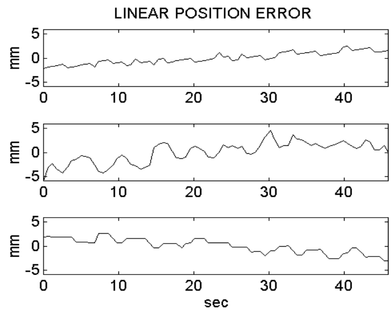


Fig. 4. Linear position error.

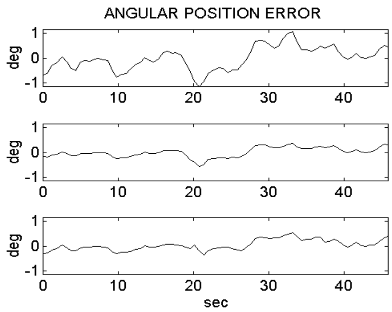


Fig. 5. Angular position error.

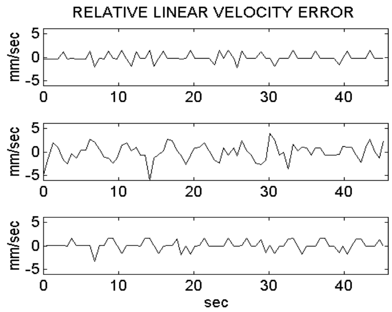


Fig. 6. Relative linear velocity error.

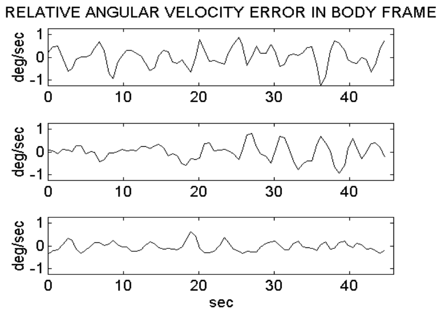


Fig. 7. Relative angular velocity error in body frame.

The reference frame is so defined that the "TV-Trackmeter" aiming direction lies along the y axis. It is worth noting the following features:

- The linear position error is about ±3 [mm] @ 3.5 [m] for the x and z axes and ±5 [mm] @ 3.5 [m] for the y axis.
- The angular position error about the y axis is the smaller one, about ±0.5 [deg], while it is ±1 [deg] on the remaining axes.

The overall result is thus more sensitive to the "TV-Trackmeter" measure errors along its aiming direction.

A pretty same result has been obtained for the linear speeds.

- The linear speed error is about ±2 [mm/s] @ 3.5 [m] for the x and z axes and ±3.5 [mm/s] @ 3.5 [m] for the y axis.

A slightly different behaviour has been obtained for the angular speed error, being ω referred to the solidal reference frame that is actually a moving frame.

- The angular speed error for every axis is mostly less than ±0.75 [deg/s], with peaks of ±1 [deg/s].

The linear position error has turned out to be fully in agreement with the error analysis of section 2.G which gives a standard deviation of 4 [mm] @ 3.5 [m] for the single point measurement.

5. Conclusions and further developments

The use of a computer stereovision and tracking system has been envisaged to allow a faster and safer capture of tumbling satellites: an already available system (the "TV-Trackmeter") has been used in a simulated environment to track a tumbling satellite mock-up. The achieved results have proved the soundness of the used approach, providing enough accuracy with respect to the requirements for such a system.
Further tests should be performed to study the behaviour of the tracking system with much longer distance, analysing the application of the same approach (or using only the tracking function without stereo measurements) to assess relative positioning even in the fly around phase. Other important tests to be carried out, in a robotic laboratory, would be the complete simulation of the capture task, using as capture device a general purpose manipulator.
To transfer this technology to space applications the next step would be the development of a space qualified system.

References

[1] J. Busing, "ROTEX - The Robotic Technology Experiment on Spacelab D2 Mission", Proc. First European In-Orbit Operations Technology Symposium, Darmstadt, W. Germany, ESA SP-272, November 1987.

[2] G. Hirzinger, B. Brunner, J. Dietrich, J. Heindl, "ROTEX - An Advanced Space Robot Technology Experiment", Proc. Third European In-Orbit Operations Technology Symposium, Noordwijk, The Netherlands, ESA WPP-059, June 1993.

[3] C. Fagerer, G. Hirzinger, "Predictive Telerobotic Concept for Grasping a Floating Object", Proc. IFAC Workshop on Spacecraft Automation and On-Board Autonomous Mission Control, Darmstadt, September 1992.

[4] W. De Peuter et al., "Satellite Servicing in GEO by Robotic Service Vehicle", ESA Bulletin, 78, May 1994.

[5] R.M. Murray, Z. Li, S.S. Sastry, *A Mathematical Introduction to Robotic Manipulation*, CRC, 1994.

[6] Hamilton W.R., *Elements of quaternions*, Chelsea, New York.

[7] W. Feller, An introduction to probability theory and its applications, New York, 1971.

[8] L. Ljung, "Asymptotic behaviour of the extended Kalman filter as a parameter estimator for linear system", IEEE Trans. Automatic Control, 1979, pp. 36-50.

[9] L. Ljung, T. Söderstrom, *Theory and practice of recursive identification*, MIT Press, 1983.

[10] Y. Bar-Shalom, T. E. Fortmann, *Tracking and Data Association*, Academic Press.