Proceeding of the 7th International Symposium on Artificial Intelligence, Robotics and Automation in Space: i-SAIRAS 2003, NARA, Japan, May 19-23, 2003

Fast Spacecraft Pose Estimation based on Zernike moments

Fernando Gandía, Alberto García Casas
GMV S.A.
Isaac Newton 11, P.T.M. Tres Cantos, 28760 Madrid
Tel: +34 91 807 21 00; Fax: +34 91 807 21 99
e-mail: fgandia@gmv.es; e-mail: albgarcia@gmv.es

Keywords Optical navigation, pose estimation, computer vision, image invariants, Zernike moments.

Abstract

Autonomous relative navigation is becoming a key feature in a variety of space vehicles and systems. A wide range of operations and activities performed by these spacecrafts demand ability to estimate relative position and attitude with respect to another satellite or space station. Furthermore, some of these activities call for additional recognition and inspection capabilities, which impose new requirements to the onboard systems. Optical based solutions allow coping with these tasks, providing a high degree of flexibility and autonomy.

This paper examines a new approach to fast pose estimation from spacecraft images, based on the computation of a kind of orthogonal image invariants: the Zernike moments. The complete development process of a fast pose estimation workbench is explained. It includes the development of an image generation tool, based on ray-tracing from a 3D model of the target spacecraft together with the pose identification algorithms.

1. Introduction

In the last years, the number of applications of fast pose estimation is growing proportionally to the complexity of the tasks envisaged for different scenarios:

- Autonomous rendezvous and docking
- In-space inspection and service activities
- In-space assembly operations

The accuracy and output frequency of the estimations required may vary for different applications but they highly influence the selection of the adequate equipment and processing technique. In fact, a compromise between accuracy and output frequency is always assumed.

New computer vision techniques coming from different application fields are contributing to improve the performances of optical navigation systems in space as well as to extend their abilities. Image invariants, such as Zernike moments, are widely used for pattern recognition in a number of common applications such as written character recognition or human face identification. The present research extends the scope of these techniques to the estimation of 3D pose from images of a spacecraft.

The main feature of the Zernike moments is their rotational invariance. This characteristic, combined with the use of normalization algorithms to achieve scale and translation invariance, makes them useful for pattern recognition. Nevertheless, they can also be applied to pose estimation whenever a 3D model of the target object is available. Zernike moments can be used as image descriptors or classifiers that allow a fast retrieving of 3D pose, given a 3D model of the target. This strategy allows reducing the amount of computational load and consequently the time inverted to provide an estimation of the target pose, by matching the vector of computed moments (feature vector) against the vectors in a pre-computed database.

Traditional pose estimation methods are usually based on high-level geometric information extracted from the target object image or from a reduced region of it, which acts as a pattern. It involves the use of different segmentation techniques that are, in general, sensible to noise and minor changes on illumination characteristics. On the other hand, methods based on low-level pictorial information present the advantage of more reduced noise sensitivity and higher robustness. The proposed method uses the low-level information provided by a set of Zernike moments computed over the illuminated pixels of the image. The adequate classification and interpolation strategies are also discussed in order to achieve the required compromise between accuracy and computation load.

The availability of “real” spacecraft images usually becomes usually a serious trouble for development of pose estimation techniques. The simulation of the spacecraft and its environment by means of scale models and robotic workbench becomes expensive and introduces additional stochastic variables to the imaging process. The development of an image generation utility seems to be an effective option to the cost of slight loss of realism. The tool developed in the frame of the present investigation is able to produce spacecraft pictures from a 3D geometric model of the target. It fulfils two different objectives:

- Generation of training images, to be used in the development and tuning of the algorithms.
- Simulation of real images for pose estimation.

Pre-processing strategies must also be carefully selected. The normalization process allows achieving near invariance characteristics with respect to translation and scale. It means isolation of pose features, discarding the influence of the relative position in the estimations.

Feature vectors containing a finite set of Zernike moments computed over normalized training images are used to create a complete description of the different possible target poses. Once a real image has been processed, its multidimensional feature vector is matched against the complete set of training vectors and mapped onto an interpolated pose. Decision-making and interpolation strategies are key elements of the complete process.

The method described could be applied at different levels of an on-board system:

- as a stand-alone system for fast pose estimation.
- as secondary or back-up solution in collaboration with a primary system.
- as a fast initial pose estimation system for providing a guess value.

2. Image Generation

As already advanced, need for an image generation utility is twofold, first the simulation of “real” spacecraft pictures taken from another spacecraft in orbit and second, the production of training images for development and tuning of pose identification algorithms.

Generation of spacecraft pictures should be kept under controlled and known conditions. It means that pictures should fit a given model of the target satellite and camera, and reproduce the effects of different factors such as relative position, attitude and camera pointing direction. Moreover, the environment characteristics (mainly lighting intensity and direction) should be considered in order to obtain enough realistic pictures. These requirements outline a globally complex task that must be handled by the generation utility.

The implementation concept selected is based on ray-tracing principles. Ray-tracing techniques try to reproduce the physical propagation of the light as a ray. The common approach for image generation consists of tracing the direction followed by light in the inverse sense (from CCD pixels to the illuminated objects in the scene and finally to the light source). In this way, the contribution of different light sources can be computed over each pixel, producing realistic images (see Figure 1).

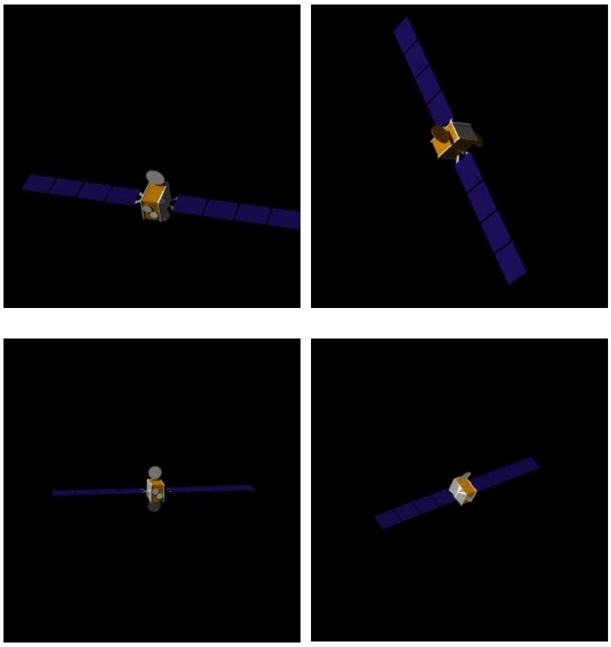

Figure 1 Spacecraft images generated at different poses

Some remarkable features of the tool are hereafter explained:

- The composition of the scene can be freely decided, including complex bodies and materials with different visual properties. The scene description is made by means of a text file written according to a natural description language convention. This file is read and interpreted by the generation routines. It includes:
    - Objects geometry definition. The spacecrafts are described as a set of basic shapes (cylinders, boxes, spheres, cones, triangular surfaces…). The basic shapes can be grouped and moved as a single object.
    - Visual characterization of the objects. A set of materials can be defined by means of their diffuse and specular RGB components. Additionally, multiple reflection and refraction characteristics can be specified. It is also possible the definition of material textures by means of texture files.
    - Position and attitude of the objects. Initial position of the basic shapes and their rotation with respect to the target reference system are described in the definition file and interpreted by the generation routines.
- The modelling of the light propagation includes diffuse, specular and ambient components.

Additional effects such as shadowing, multiple reflection and refraction are also computed by the generation routines.

- The time evolution of the vehicle position and attitude is specified for each different picture. A rotation matrix is defined for the whole object to be moved, together with the camera position with respect to it and its pointing direction.
- The image is generated by computation of the intersection points between light rays and objects, providing the total intensity of each pixel. The free code Raytressi is used for this purpose. Anti-aliasing algorithms are included in the computation in order to avoid non-desirable effects of pixel quantization. The lighting model used, includes both specular component and diffuse component. Additionally a very small ambient component is included to approximate the light diffusely reflected from other surfaces or coming from sources that are difficult to characterise (stars background…):
    - The diffuse component considers the surfaces following the Lambert´s law. It means that the light reflected from a surface in a determined direction is proportional to the cosine of the angle between the surface normal and that direction.
    - The specular component approximates the behaviour of shiny surfaces. If a surface is an ideal mirror, the light from a source will reach the CCD bouncing of a fixed point on the surface only if the direction to the light coincides with the reflected direction to the sensor.

The equation for total illumination contribution over a pixel is:

$$I_{total} = k_{amb} I_{amb} + \sum_{lights} I_i (k_{diff} (\vec{L}_i \cdot \vec{N}) + k_{spec} (\vec{L}_i \cdot \vec{R})^p)$$

Equation 1: Lighting model

Where:

I_{total}	is the total light intensity over a pixel
I_{amb}	is the ambient light intensity
I_i	is the intensity of the i-th light source
k_{amb}	is the ambient reflection coefficient of the surface
k_{diff}	is the diffuse reflection coefficient of the surface
k_{spec}	is the specular reflection coefficient of the surface
$\vec{L}_i$	is the direction to the i-th light source
$\vec{N}$	is the direction of the surface normal
$\vec{R}$	is the reflected direction to the sensor
p	is the phong coefficient

3. Image Processing

Pose estimation techniques commonly used till now focus on tracking of a predefined pattern or set of landmarks. The troubles associated to this approach are related to the lack of flexibility (need for continuous visibility of the pattern or need for human intervention for selection of landmarks) together with high noise sensibility levels. The strategy proposed (based on low level pictorial information) allows increasing the level of autonomy and robustness by neglecting high-level information, whose extraction involves more complex and perturbation-sensitive algorithms. The developed processing technique follows a series of sequential steps:

- Pre-processing. Its purpose is the minimisation of the effect of all kind of factors but the vehicle attitude. It means achieving invariance with respect to the spacecraft relative position. The resulting image is, then, ready for further steps.

- Zernike moments computation. The image is numerically characterised by means of a series of invariants: the Zernike moments, computed from low-level pictorial information.

- Pose retrieving. The feature vector of Zernike moments belonging to the target picture is matched against a set of feature vectors in a database. The pose estimation is retrieved from the database.

3.1 Image pre-processing

Pre-processing routines implemented are devoted to picture conditioning for further processing. The algorithms applied on input images try to minimise, in this case, the effects of changing environment conditions and other variables different from the desired one: the target vehicle relative attitude.

A first strategy, able to reduce the variability introduced by rapidly varying illumination conditions, consists of using binary images for all subsequent computations. It seems a good approach whenever the complete spacecraft pattern could be highlighted against an empty background (quite feasible at certain conditions for space applications). In this way, the impact of differences in pixel intensity for images containing the same target pose at different light conditions is avoided. A standard pixel threshold algorithm is used for obtaining these binary images. Pixels above a constant value are set to white while pixels below it are set to black.

Once again, the goal is the isolation of the spacecraft pose information, independently from other variables. The relative position of the target with respect to the camera determines the x-y position in the picture and the scale factor or size. As we will show in section 3.2, these two features have important effects on the values obtained for Zernike moments, used as image descriptors. Therefore, a normalization algorithm has to be applied on binary images previous to Zernike computation.

Normalization strategies widely employed in different application fields use the zero-th order geometric moment (area of illuminated pixels in the binary image) for scaling the image objects. However, this strategy seems not to be the most adequate for target images varying significantly in total pictorial size. The approach selected uses the concept of the "bounding volume" of the target object for scale normalization. This process can be described as follows:

- A rectangular bounding volume of the target object is computed on the binary image.
- The subset of pixels contained inside the bounding volume is resized to a constant dimension.
- The centroid of the target object is computed.
- Finally, the resized bounding volume image is placed inside an empty picture such that the centroid of the object is at the geometric centre.

As observed in Figure 2, the use of a normalization algorithm achieves near-equivalent binary images for an original scene showing the same target attitude but different scale and position in the image plane.

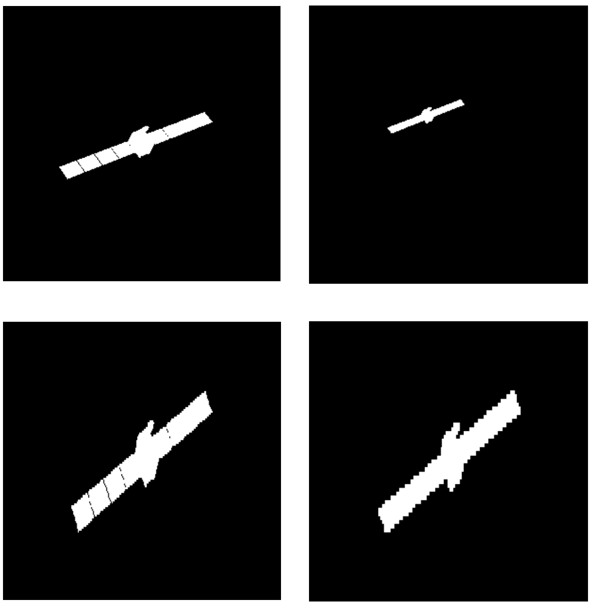

Figure 2 Effects of the normalization process (scale invariance)

3.2 Image descriptors. The Zernike moments

Once the spacecraft attitude information has been isolated from the remaining contributions the following problem is to interpret this information in order to obtain a reliable pose estimation.

Our proposed techniques focus on the characterisation of contents of each picture by means of a set of image descriptors. At this step, the characterisation of the image pictorial contents is equivalent to the characterisation of the spacecraft pose (due to minimisation of other effects by using convenient pre-processing algorithms). Therefore, a pose estimation can be derived by matching against a previously computed correspondence table of image invariant vectors and spacecraft attitudes.

The image invariants computed for each picture are the Zernike moments. These moments have been traditionally used as invariant global features of a picture in different applications regarding to pattern recognition and image classification. The Zernike moments present the advantage of completeness and orthogonality, which allow representing any square integrable function defined on the unit disk. Furthermore, their rotational invariance (in module, and only approximately for digital images) makes them extremely useful in a number of applications where rotation in the image plane must be ignored.

The formulation implemented for discrete computation of Zernike moments of p-order and repetition q is the following:

$$Z_{pq} = \sum_x \sum_y I(x,y) \cdot V^*_{pq}(\rho,\theta)$$

$$V_{pq}(\rho,\theta) = e^{jq\theta} \sum_{l=0}^{(p-|q|)/2} (-1)^l \frac{(p-l)!}{l!\left(\frac{p+|q|}{2}-l\right)!\left(\frac{p-|q|}{2}-l\right)!} \rho^{p-2l}$$

Equation 2: Zernike moments

Where:

p	is the moment order (non-negative integer)
q	is the moment repetition, subject to $p-\|q\| = even$
$I(x,y)$	is the intensity for the pixel at coordinates x,y
ρ	is the length of the vector from image centre to pixel coordinates x,y $\rho = \sqrt{x^2+y^2}$, where $x^2+y^2 \le 1$
θ	is the angle between vector ρ and the horizontal axis; $\theta = \arctan(y/x)$

Completeness and orthogonality make Zernike moments especially suitable for image classification (avoiding information redundancy) and reconstruction. Equation 3 allows recovering an image from an infinite series of Zernike moments:

$$I(x,y)=\sum_{p=0}^{\infty}\sum_{q=-p}^{p}\tau_p Z_{pq}V_{pq}(x,y)$$

with $\tau_p = \dfrac{p+1}{\pi}$

Equation 3: Image reconstruction from Zernike moments

Figure 3 shows an example of image reconstruction from moments till order 20 for a normalized 100x100 pixel picture.

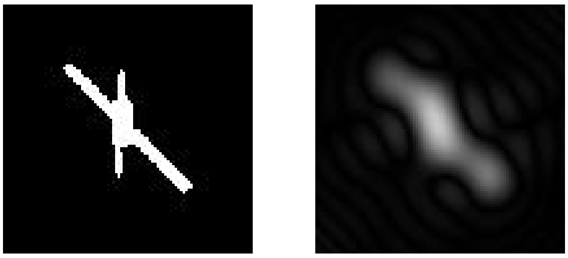

Figure 3 Image reconstruction from moments till order 20

In our case, orthogonality and completeness of Zernike moments allow using them as efficient image descriptors.

A compromise between computation load and accuracy in the characterisation of the pictorial information has been assumed. In fact, the implemented algorithm computes the first 66 moments (till order 10), which seems to be enough in terms of accuracy while maintaining low computation load.

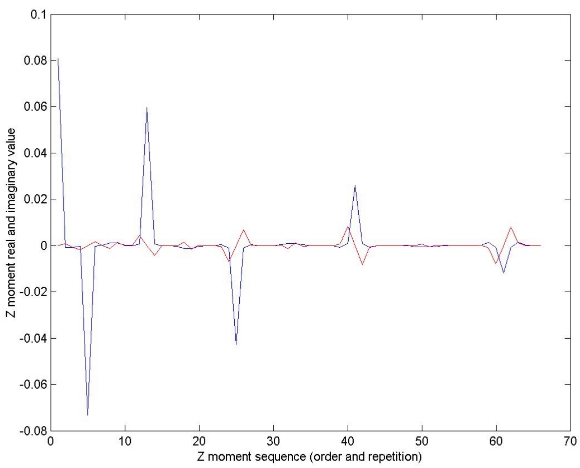

Figure 4 Zernike moments sequence till order 10

3.3 Pose estimation

The philosophy of the pose retrieving mechanism focuses on the pre-existence of a correspondence table for feature vectors (Zernike moments vectors) and spacecraft relative attitudes (Euler angles). It makes the pose retrieving technique dependent on the target spacecraft model, requiring different tables for different spacecraft configurations.

Known the feature vector of a "real" spacecraft image, it can be compared to a set of feature vectors stored in the correspondence table. At this time, the attitude values corresponding to the closest feature vector are selected.

Different options were considered for classification or matching of the feature vector (neural networks, spherical manifolds, use of different error indices…). The matching method finally selected entails a quite simple but robust approximation to the problem. It consists on the computation of a simple index error for each comparison between the "real" feature vector and a feature vector in the table. This index error is calculated as the sum of the squares of the differences of all Zernike moment components. Then, the minimum of the index error is searched among all comparisons performed. For a correspondence table characterised by a fine-enough generation step, the pose estimation given by the minimum should provide values reasonably close to the real ones.

Additionally, the error values corresponding to the neighbour points of the table (in terms of Euler angles), together with the knowledge that the real minimum of the error should be zero are used to improve the estimation.

The correspondence table has been generated with the first 66 moments for a 20 degrees equi-spaced matrix of all possible spacecraft attitudes. It means that a total amount of 3640 training images have been produced and processed to conform the table.

4. Analysis of Results

Preliminary tests performed on single modules of the workbench have demonstrated their ability to proceed with further analyses and tuning activities. These tasks should permit to assess the feasibility of the techniques proposed.

First cases were run with a simplified satellite model, characterised by a high degree of symmetry. It showed that pictorial ambiguity between different poses causes difficulties for identification of the correct spacecraft attitude in many cases (see Figure 5). Nevertheless, miss-matching of spacecraft poses presents an interesting feature: principal directions of the target in terms of "mass" distribution (e.g. solar arrays principal axis) are successfully identified. It has valuable consequences in terms of safety due to the fact that position of spacecraft elements with larger collision risk is always properly known.

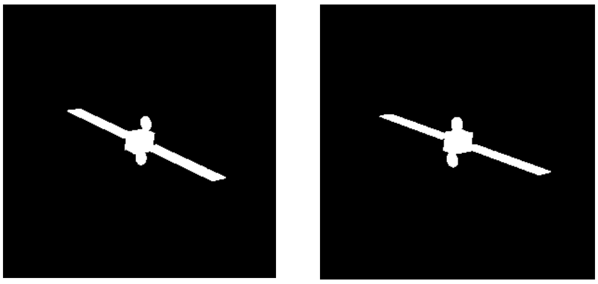

Figure 5 Retrieving errors due to model symmetry

A decrease of the database generation step size and/or an increase in the dimensions of the binary normalized images should likely improve the classification performances for highly symmetric models.

Analyses performed on more asymmetric targets confirmed that miss-matching was principally due to pictorial similitude in the image plane caused by model ambiguity for different poses. It also showed that it is a phenomenon difficult to avoid in many cases.

Figure 6, Figure 7 and Figure 8 show roll, pitch and yaw errors (and mean values) achieved for 100 random samples using 60x60 pixels normalization and data-base of 3640 images.

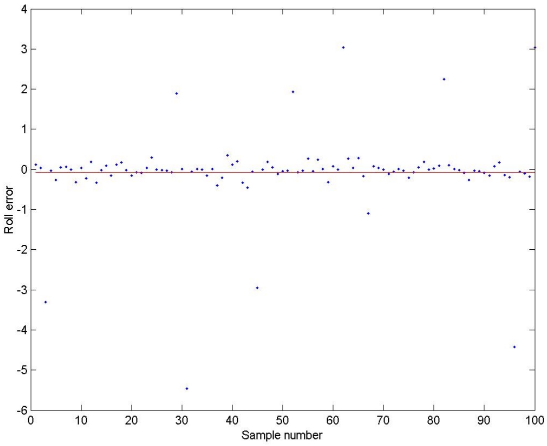

Figure 6 Roll identification error

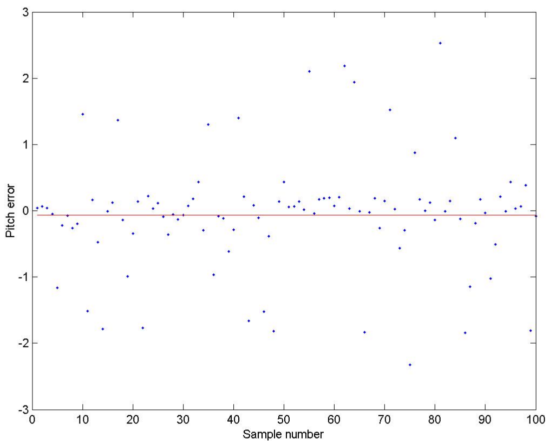

Figure 7 Pitch identification error

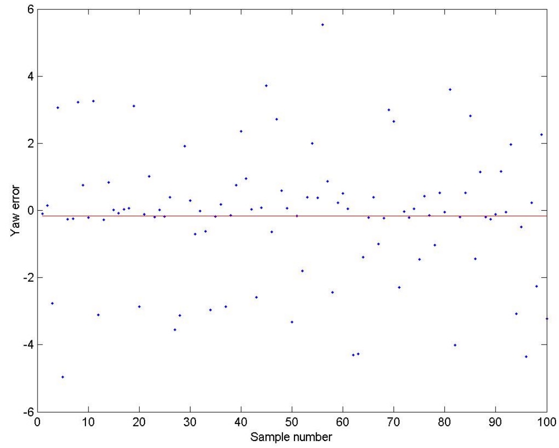

Figure 8 Yaw identification error

The analysis of retrieving errors for roll, pitch and yaw rotations reveal that exist privileged axes determined by the distribution of the spacecraft major elements. In this case, roll retrieving errors are considerably lower than pitch or yaw errors due to change of solar array principal axis direction when changing roll angle. Figure 9, Figure 10 and Figure 11 allow a better understanding of this observation. It can be noted that a larger amount of samples with error under the size of the database generation step (represented by a line in the figures) is achieved for roll and pitch with respect to yaw.

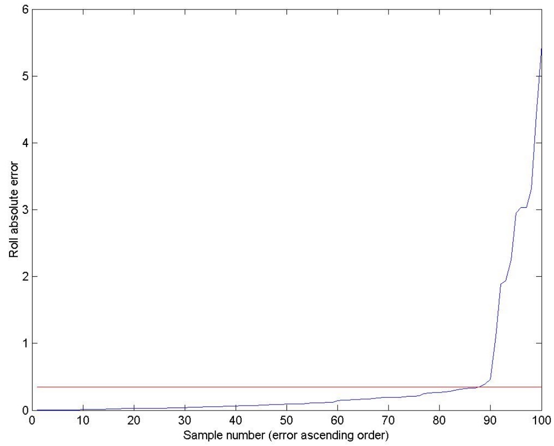

Figure 9 Roll absolute error distribution

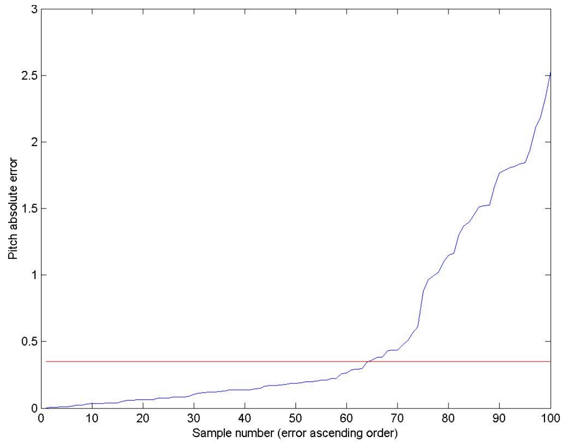

Figure 10 Pitch absolute error distribution

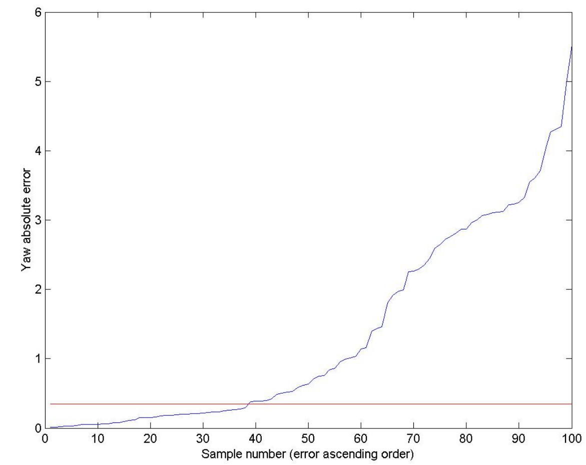

Figure 11 Yaw absolute error distribution

The characterisation of the absolute error distribution for correctly matched samples is presented in Table 1.

Abs. error	Roll	Pitch	Yaw
Mean	0.1071	0.1217	0.1532
Std. Dev.	0.0972	0.0843	0.0995

Table 1 Absolute error distribution for correctly matched samples

Despite the increment of identification performances for more asymmetric targets, visual ambiguity remains in some cases and it is difficult to completely avoid it (see Figure 12). The generation of a larger database of Zernike moments should minimise this effect.

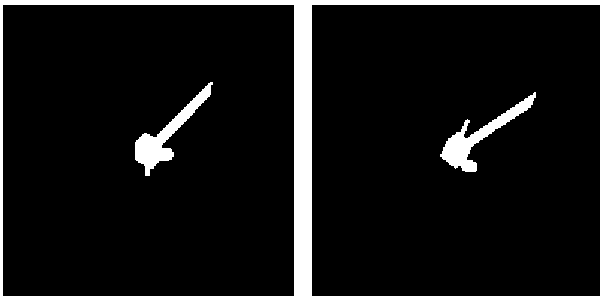

Figure 12 Retrieving errors due to visual ambiguity

Invariance of performances with respect to scale and translation in the image plane has also been tested, demonstrating optimal behaviour of the algorithms.

Figure 13, Figure 14 and Figure 15 show minimum change on identification performances for a scale change of 10% and translation inside the image plane.

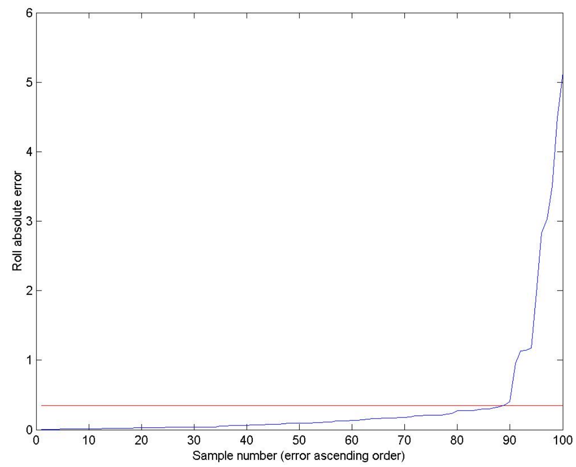

Figure 13 Roll absolute error distribution for a scaled sample

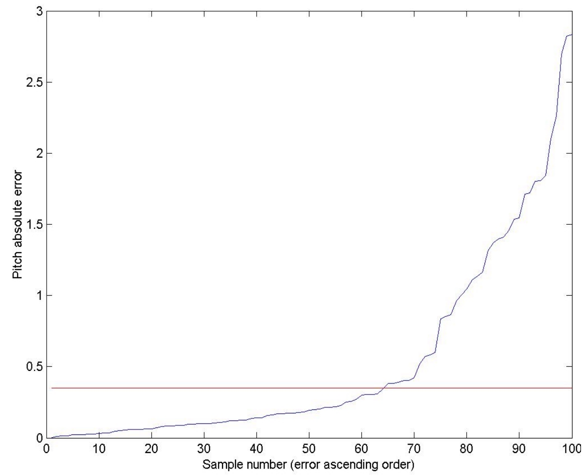

Figure 14 Pitch absolute error distribution for a scaled sample

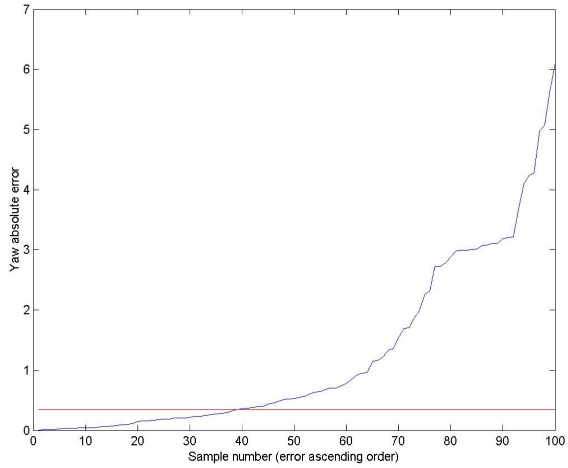

Figure 15 Yaw absolute error distribution for a scaled sample

5. Conclusion

A novel approach to fast pose estimation has been proposed, based on the computation and comparison of a series of orthogonal image descriptors, the Zernike moments. The development of image generation and processing utilities has allowed assessing the feasibility of using these new concepts for spacecraft pose retrieving. Regarding to major characteristics of the technique proposed, the following conclusions can be highlighted:

- The proposed processing technique, based on low-level pictorial information, provides a higher degree of flexibility (no need for pattern visibility or human intervention for landmark selection) than conventional methods. Moreover, the robustness against changing illumination conditions, noise and target position considerably increases.
- Computing load is reduced significantly due to the use of a reduced processing chain and simple computational algorithms.
- Pose estimation performances are considerably influenced by the characteristics of the target object (ambiguity due to geometrical symmetry…) and the size of the correspondence database.
- The proposed method works in a more human-like way than traditional pose estimation approaches. It makes it to inherit some of the characteristics of human vision: great flexibility but possible ambiguity troubles together with medium accuracy level. These characteristics make it highly suitable for use in combination with traditional systems, providing system initialisation values or backup measurements. Further refining of design parameters should be desirable for improvement of classification performances.

References

[1] S.X. Liao, M. Pawlak. *A Study of Zernike Moment Computing.* Proceedings of the Asian Conference on Computer Vision ACCV 1998.

[2] K. Chang, J. Ghosh. *Three-Dimensional Model-based Object Recognition and Pose Estimation using Probabilistic Principal Surfaces.* Proceedings of SPIE 2000.

[3] D. de Ridder, O. Lemmers, R.P.W. Duin, J. Kittler. *Image Database Retrieval using Adaptative Non-linear Image Descriptions.* Proceedings of the Advanced School for Computing and Imaging ASCI 2000.

[4] S.X. Liao, M. Pawlak. *On the Accuracy of Zernike Moments for image Analysis.* IEEE Transactions on Pattern Analysis and Machine Intelligence 1998.

[5] P.C. Hew, M.D. Alder. *Zernike or Orthogonal Fourier-Mellon Moments for Representing and Recognising Printed Digits.* University of Western Australia. 1998

[6] C.Chee-Way, R. Mukundan, P. Raveendran. *An Efficient Algorithm for Fast Computation of Pseudo-Zernike Moments.* Proceedings of the Image and Vision Computing New Zealand IVCNZ 2001.