Image-based aircraft pose estimation using moment invariants

Marcel G. Breuers

TNO Physics and Electronics Laboratory,
P.O. Box 96864, 2509 JG The Hague, The Netherlands

ABSTRACT

The problem of estimating aircraft pose information from mono-ocular image data is considered using two different pose estimation algorithms. Both algorithms are based on the rotation invariant moment approach that was introduced by Dudani. The dependence of pose estimation accuracy on image resolution and aspect angle was investigated through simulations using sets of synthetic aircraft images. It is shown that increased pose-estimation accuracy can be obtained by breaking the nearest neighbour search procedure in two parts.

Keywords: Moment invariants, pose estimation, computer vision

1. Introduction

In order to use air defence artillery effectively against incoming aircraft, it is necessary to make accurate predictions of aircraft trajectories. Traditionally, systems capable of pointing and tracking use only position measurements based on radar data to update trajectory estimates. The performance of these systems becomes poor when confronted with fast manoeuvring targets. It is known from literature[1,2] that fusion of target *position* data gathered by radar and target *pose* estimates extracted from camera images results in improved tracking and prediction performance. The fusion process is shown in Figure 1.

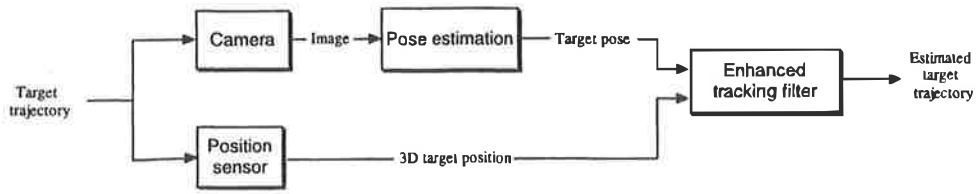

Figure 1: Pose estimation as a part of an enhanced target tracking system

In this paper we discuss the design, implementation, and testing of an algorithm capable of extracting aircraft pose information from camera images. First, an introduction is given to pose estimation techniques and results from a literature search are presented. Next, the details of the implemented pose estimation algorithm are explained. Finally, the results from numerical experiments are discussed.

2. ANALYSIS

The general problem of automatic image based object recognition and pose estimation has been frequently studied in the field of computer vision and pattern recognition. Some of the techniques developed in this field are particularly suited for automatic aircraft pose estimation. This section will present a short outline of general pose estimation techniques, taking into account the advantages and disadvantages of mono-ocular and stereo images, the choice between local and global image features, and the way in which these features can be converted to a pose estimate.

2.1 Image based pose estimation

At the start of the information processing chain, a choice has to be made between the use of 3D information from stereo image pairs and the use of 2D information from mono-ocular images. Accurate 3D shape information can help to reduce object pose ambiguity. However, extracting accurate 3D information from stereo images is only possible if the distance

Part of the SPIE Conference on Automatic Target Recognition IX • Orlando, Florida • April 1999
SPIE Vol. 3718 • 0277-786X/99/$10.00

applied. In the first part of the algorithm, a watershed algorithm is used to separate object and background. Multiple separation hypotheses are generated, converted to Fourier descriptors, and compared to a library by means of nearest neighbour search. If no match is found during the search, it is assumed that the object separation was not successful. In this case, local image features and syntactic pattern matching are applied. Details of this approach are not provided in their paper. To further enhance the robustness of object recognition, evidence is accumulated over time. An estimate of the pose recognition reliability is computed on line. It is assumed that the pose ambiguity problem can be solved by using trajectory information. The performance of the system was investigated using simulations. During these simulations, parameters such as background complexity, contrast, resolution, number of target types, and object pose were taken into consideration. In addition, they examined the sensitivity of the system performance to target shape modifications caused by presence or absence of ordinance systems. The overall conclusion is that good pose estimation and recognition performance can be achieved. No quantitative figures are provided.

3. The Pose estimation algorithm

The results from the literature study that were discussed in the introduction make it clear that most methods of image based aircraft pose estimation are based on the same strategy. In the first step of this strategy, rotation, translation, and scale invariant image features are used to compute the aircraft rotation about two axes that are perpendicular to the optical axis of the camera. In the second step, translation and scale invariant image features are used to compute the aircraft rotation about the optical axis of the camera. This two step strategy is the basis of our pose estimation algorithm.

The algorithm's performance is substantially influenced by the image feature type that is used. Most methods known from literature rely on moments or Fourier descriptors. A comparison of the aircraft *classification* accuracy that can be obtained by various feature types was made by Reeves[6]. His experiments indicate that Fourier descriptors result in a classification accuracy of approximately 90%, while moment invariants give approximately 65% *classification* accuracy. However, Dudani[4] and Wintz[7] claim for these two feature types a *pose estimation* accuracy that is very similar. In our algorithm, moment invariants are used as they can be easily implemented and are likely to give good performance.

In most aircraft pose estimation algorithms the input image features are mapped to a pose estimated by means of a nearest neighbour search procedure. The features can be compared either directly in the original feature space, or they can be transformed first to a new feature space. Two arguments are used in favour of the second option: better classification accuracy and reduction of the feature space dimensionality. The latter results in reduced computational complexity. Reeves[6] argues that classical feature vector conditioning procedures, such as the eigenvalue transformation, do not necessarily result in performance improvement since the within-class feature vector variation is, in general, greater than the between-class feature vector variation. Reeves therefore uses a "variance balancing" technique to condition the features vectors before they are compared. The purpose of this technique is to make the effect of the various feature vector elements more equal. His experiments show that this technique works well provided that higher order moments, which are very noise sensitive, are excluded from the feature vectors.

The structure of the implemented pose estimation algorithm is visualised in Figure 2. The processing steps that take place in the conversion from input image to pose estimate are described below:

1) The grayscale image from an infrared camera is converted to a binary image by means of thresholding. Most of the gray level information within an object changes considerably depending on environmental conditions and can therefore be ignored. A noticeable exception to this general rule is the exhaust outlet, since it contains relevant pose information and can be easily discerned in a grayscale image. It is assumed that the aircraft silhouette contains sufficient information for object pose estimation. In addition, we assume that only one target is visible in the image and that sufficient contrast between target and background is available to allow for simple and robust image segmentation.

2) From the *binary* silhouette image, a vector, $\underline{C}_{in}$ of central scale invariant moments is computed. In later processing steps the central moments are used to derive Hu-moments and to compute the aircraft rotation about the optical axis of the camera.

4) Scaling and offset are applied to each element of the Hu-moment vector $\underline{H}_{in}$. This is necessary since the absolute value of the higher order Hu-moments is in general several orders of magnitude smaller than the absolute value of lower order Hu-moments. The scaling and offset values are based on the variance and mean of the Hu-moments in the reference database. As a result of this procedure, zero mean and unit variance are obtained for each feature vector component when averaged over the complete training set.

5) The normalised Hu-moment vector, $\underline{H}_n$, is compared to the reference database of Hu-moment vectors, using the Euclidean distance as metric. The result from this search procedure is the best matching Hu-moment vector in the database, denoted $\underline{H}_m$. The corresponding central moment vector in the database is denoted $\underline{C}_m$, the corresponding aircraft azimuth and elevation angles in the database are denoted α, and β.

6) The central moment vector $\underline{C}_{in}$ is compared to $\underline{C}_m$ to find the aircraft rotation γ about the optical axis of camera. The rotation angle γ is computed using the equation,

$$\gamma = \frac{1}{2}\arccos\left((M_{20} - M_{02})\cdot(\tilde{M}_{20} - \tilde{M}_{02}) + 4M_{11}\cdot\tilde{M}_{11}\right),$$

where M_{20}, M_{02}, M_{11} and $\tilde{M}_{20}, \tilde{M}_{02}, \tilde{M}_{11}$ are respectively central scale normalised moments moments from the input silhouet image and the corresponding reference set image. This procedure cannot distinguish between two object rotations that differ a multiple of 180°. This problem is dealt with in step 7.

7) From the basic pose estimate [α, β, γ] that was derived thus far, a number of alternative pose hypotheses are derived to account for the ambiguity that is induced by the object symmetry, the mapping from a 3D object to a 2D image and the limitations in the γ-estimation procedure. The causes of pose ambiguity and the related pose equivalence relations are summarised in Table 1. The aircraft pose hypotheses that are generated in processing step 7 are listed in Table 2.

Table 1: Three causes of object pose ambiguity

Equivalence relation	Cause
[α,β,γ] ↔[-α,-β,γ-180°]	Rotation symmetry (180°) relative to aircraft main body axis
[α,β,γ] ↔ [180°-α,-β,γ]	Mirror symmetry relative to the plane through main body axis and tail wing
[α,β,γ] ↔ [α,β,γ+180°]	Limitations in procedure that estimates aircraft rotations relative to the camera axis

Table 2: Eight aircraft pose hypotheses, based on the initial pose estimate [α, β,γ]

azimuth	elevation	roll
α	β	γ
180°-α	-β	γ
180°-α	β-180°	γ
α	-β-180°	γ
α	β	γ+180°
180°-α	-β	γ+180°
180°-α	β-180°	γ+180°
α	-β-180°	γ+180°

8) The eight hypotheses that were generated in step 7 are compared to the previous pose estimate, using the pose similarity measure described in paragraph 4.1. The current pose estimate will be set equal to the best matching hypothesis.

4.2 Pose estimation accuracy and aspect angle

Many aircraft pose estimation methods known from literature[3,4] are based on the assumption that the symmetry properties of an aircraft allow for omitting a considerable number of aspect angles from the reference database without much degradation of estimation accuracy. In this section, we investigate the validity of this assumption by analysing data from numerical experiments based on reference feature vector sets covering two different ranges of aspect angles. In addition we examine the dependence of the estimation accuracy on the aspect angle of the input images.

In the first experiment the pose estimation algorithm had access to a limited set of reference features vectors covering azimuth angles in the range 0° through 90° and elevation angles in the range -90° through 90° (Figure 5). A test set of images showing an NF5 aircraft from aspect angles with azimuth in the range 0° through 360° and elevation in the range -90° through 90° was processed by the algorithm. The roll angle of the aircraft in the input image data set was 80°. The resolution of both the images in the test set and the reference set was 100x100 pixels. The pose estimation errors that occurred for different values of azimuth and elevation are shown in Figure 6.

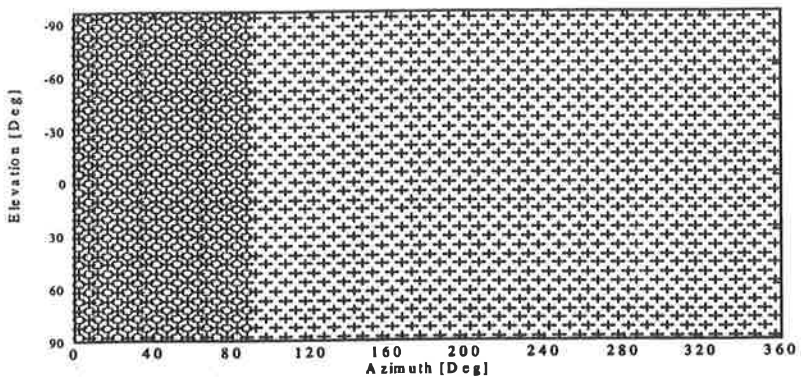

Figure 5: Overview of aspect angles in algorithm database (o), and aspect angles in test data set (+)

In figure 6 two symmetric patterns can be recognised: the first pattern occurs in the left half plane and is symmetric relative to the point were the azimuth is 90° and the elevation is 0°. The second pattern occurs in the right half plane and is symmetric relative to the point were the azimuth is 270° and the elevation is 0°. The symmetry of the pattern that occurs in each half plane is caused by the fact that the aircraft is mirror symmetric relative to a

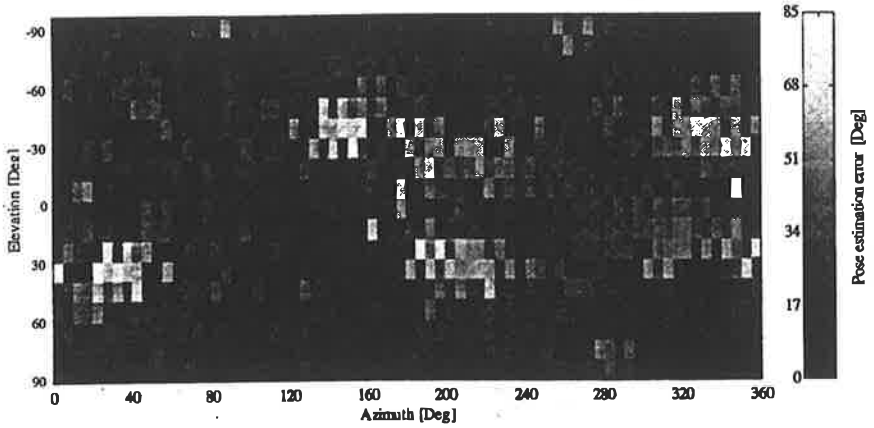

Figure 6: Pose estimation error depending on input image aspect angle. Reference database with azimuth in the range 0-90°. Test images are rotated 80° clockwise in the image plane relative to reference images. Feature vectors are based on 7 Hu-moments.

4.3 Pose estimation accuracy and target resolution

In order to find the relation between image resolution and pose estimation accuracy a number of simulations were carried out at various resolutions. Since in all these experiments aircraft images were evaluated independently, the pose ambiguity resolving part of the algorithm could not rely on previous pose estimates to select the best pose hypothesis. To circumvent this problem the eight pose hypotheses generated in step 7 of the algorithm were compared to the true aircraft pose, instead of to the previous pose estimate. In all experiments the algorithm had access to a complete database of reference moment vectors covering azimuth in the range 0° trough 360° and elevation in the range -90° trough 90°. Both the test and the reference images were taken at 100x100 pixels resolution. The test set images were rotated 80° counter clockwise relative to the images in the source data set. Figures 8 through 13 show typical NF5 images from the test data set at six different resolutions.

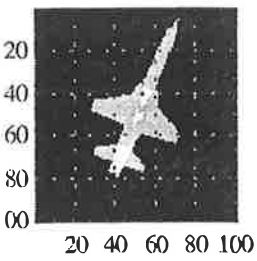

Figure 8: NF5 at 100 % resolution

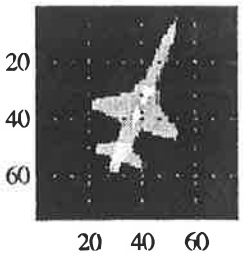

Figure 9: NF5 at 75 % resolution

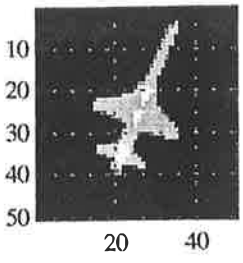

Figure 10: NF5 at 50 % resolution

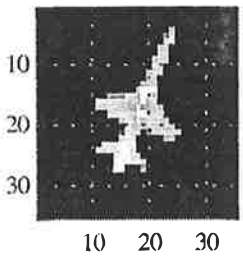

Figure 11: NF5 at 35 % resolution

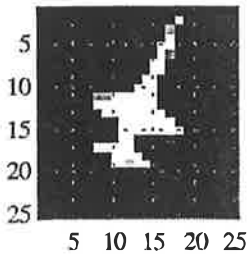

Figure 12: NF5 at 25 % resolution

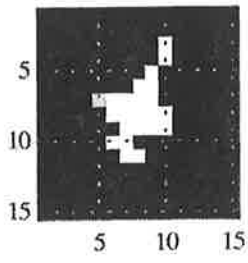

Figure 13: NF5 at 15 % resolution

The average pose estimation errors at 5 different resolutions are summarised in Table 5. These data show that the estimation performance does not change much for input images with resolution in the range 100% to 35 %. Only at 15% image resolution, the pose estimation errors become much larger.

Table 5: Average pose estimation error at various image resolutions

Relative image resolution	100 %	75 %	50 %	35 %	15 %
Average pose estimation error	14.07°	14.55°	14.21°	15.35°	23.53°
Median pose estimation error	8.77°	9.95°	10.04°	11.64°	21.26°

5. Improvements in the pose estimation algorithm

The experiments that were done to study the relation between target aspect angle and pose estimation accuracy indicate that it is advantageous to use in the first part of the pose estimation process features that are less sensitive to noise and object rotations in the image plane. The seventh Hu-moment is particularly sensitive to noise and should therefore, if possible, be omitted from the nearest neighbour search procedure. However, from the set of seven Hu-moments, only the seventh one can be used to discern between mirrored silhouettes. A number of modifications can be made to the algorithm described in

6. Summary and conclusions

An algorithm has been designed and implemented to study the performance that can be achieved in automatic image based aircraft pose estimation. Experiments show that the pose estimation procedure works fairly well down to low resolution images of 35x35 pixels, giving a median pose estimation accuracy better than 12°. It was also found that the pose estimation error was nearly independent of the input image resolution over a wide range of settings. A modification of the original pose estimation algorithm has been implemented and resulted in improved pose estimation performance.

ACKNOWLEDGEMENTS

The author is grateful for the fruitful discussions with N. de Reus, P. Ockeloen, K. Schutte, B. van den Broek, and J. Schavemaker.

REFERENCES

1. D. Sworder and R. Hutchens, "Maneuver estimation using measurements of orientation", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 4, July 1990.
2. J.D. Kendrick, *Estimation of aircraft target motion using pattern recognition orientation measurements*, Ph.D. dissertation, Air force Inst. Technology, Wright-Patterson Air Force, OH, 1978.
3. Z. Chen and S. Ho, "Computer vision for robust 3D aircraft recognition with fast library search", *Pattern recognition*, vol 24, no. 5, pp 375-390, 1991.
4. S.A. Dudani, K.J. Breeding and R.B. McGhee, "Aircraft identification by moment invariants", *IEEE Trans. Comput.* C-26, pp. 39-46, 1977.
5. T. Glais, A. Ayoun, "Image-based air target identification", *Proc. of SPIE*, vol 2298, pp. 540-551, 1994.
6. A.P. Reeves, R.J. Prokop, S.E. Andrews and Kuhl F.P., "Three dimensional shape analysis using moments and Fourier descriptors", *Proc. 7th Int. Conf. Pattern Recognition*, pp. 447-450, 1984.
7. T.P. Wallace and P. Wintz, "An efficient three-dimensional aircraft recognition algorithm using normalised Fourier descriptors", *Comput. Graphics Image Process*, vol. 13, pp. 99-126, 1980.