Aircraft Component Detection Based on 3D Object Recognition and Relative Position Estimation

Lin Tian
Instrumentation Science and Opto-electronics Engineering, Beihang University, Beijing, 100191, P. R. China
Email: lintian@aspe.buaa.edu.cn

***Abstract*—An aircraft component detection method based on 3D recognition and relative position estimation is proposed in this paper. Direct detection of components that are small parts of an object is difficult for the lack of distinctive features. Since relative position of a component to the main axis of the plane is invariant to 3D transformation, major direction vector is proposed to find search region that encloses interesting parts. Major direction vector is parallel to the projection of the main axis of a plane in 2D images. 3D recognition based on shape features is applied to estimate pose of a plane. Fourier Descriptors are applied to extract features. The detection in an image is reduced to the search region after the two steps. A detection rate of 84% is achieved in the search of landing gear.**

***Index Terms*—aircraft component detection, major direction vector, 3D object recognition.**

I. Introduction

Aircraft recognition through combined image processing algorithms is an extensively studied topic [1]-[5]. Most commonly used classifiers are neural network, SVM and Bayes classifier. For image features used for recognition, [1], [3] and [4] extract shape feature by Invariant Moments and Fourier Descriptors. However, no specific literature deals with the detection of key components on airplanes. For instance, landing gear is an important part on aircrafts. If it is not properly positioned and locked, forced landing would take place, which endangers people aboard and delays airlines. Usually pilot makes judgments according to monitor in the cabin. The detection method in this paper through 3D object recognition and relative position estimation could assist the electronic system. It could also serve as reference for ground service to coordinate landing. Detection through image processing improves efficiency compared to eye detection and reduces unnecessary communication between pilot and ground.

The detection algorithm is divided into two stages: 3D plane recognition and component search region estimation. There are two common methods for 3D recognition. Reference [6] uses Structure from Motion for scene recovery and camera motion estimation. SfM is based on 3D structure reconstruction and 2D-3D correspondence to estimate pose of 3D object in 2D testing images. The most commonly used local features during reconstruction and matching are SIFT and Harris. In [7], SfM is improved and applied to recognition and pose estimation of daily articles. The recognition rate of the improved method is between 88% and 92%. The other method is a bottom-up recognition from 2D views. Objects are recognized by matching testing images to individual 2D training images. An aspect graph is represented by prototypical views in [8]. A group of views are selected as “aspects” on the viewing sphere. A graph structure is formed where each node represents an aspect. The assumption is that arbitrary views within a range of an aspect on the viewing sphere are equivalent to the aspect. Shape similarity metric is used to match views to stored aspects. Both identity and pose are determined.

As input to classifiers, object features are extracted of the interest region. Two major categories of features are commonly used in aircraft recognition, local features and shape or structure features. Similar to recognition through 3D model, [13] extracts SIFT features from sample images. Instead of building 3D point cloud, local features are first grouped to represent parts of object according to consistency in appearance and geometry. Recognition and pose estimation are accomplished by part matching. In [14], 3D skeletons of planes are applied for recognition. The center of inscribed ball draws a course as the ball moves around in the 3D plane model. The 3D course or skeleton of the plane forms 2D projections from multiple views. Those 2D skeletons are plane image features. Shape features, such as Fourier Descriptors and Invariant Moments, are also used in plane recognition [1], [15]. In 3D recognition problem in this paper, Fourier Descriptors are used to represent shape features of arbitrary views of plane image.

The discussion is organized as follows. For recognition, Section II *A* constructs a multi-view database of four plane models. Fourier Descriptors [11], [12] are used to extract shape features during 3D recognition, described in Section II *B*. In Section II *C*, major direction vector representing the scale and relative position of search region is used to locate polygon search region in testing image. The

Manuscript received November 18, 2013; revised August 1, 2014.

doi: 10.12720/lnit.2.2.163-167

component detection task is reduced to within the search region. Experimental results are presented in Section III. The discussion is concluded in Section IV.

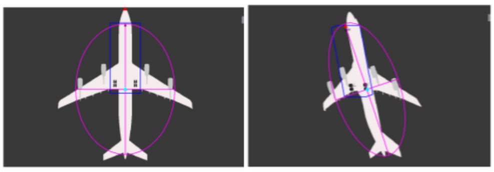
II. AIRCRAFT COMPONENT DETECTION ALGORITHM

It is difficult to reconstruct 3D model if local feature of the object is obscure or ratio of mismatches is high. In real problem, geometrical features of different planes are very different, whereas local features are ambiguous for recognition. For instance, A380 planes belonging to different companies wear their own logo and are in various colors and textures. The popular local features such as SIFT [9] and Harris [10] fail in recognition in this problem. 3D object recognition method based on a collection of 2D views of aircraft is used in this paper.

Plane recognition is the base for search region estimation. The pose of matched object in database is taken as the pose of the testing object. Relative position of the search region in testing image is determined by planar transformation between the best match and testing image. The major direction vector representing scale and rotation angle is applied to compute the 2D transformation. The whole database includes 2D shape features of prototypical views, search region and major direction vector corresponding to each view.

A. Sample Image Acquisition and 2D Representation of Search Region

3DMAX is used to generate sample images with known object angle and coordinates of interesting parts. Prototypical views of plane model are extracted with an interval of 15 degrees. The view space is divided into $24 \times 24 = 576$ sub-regions. Since the structure of plane is symmetrical, only a quarter of the viewing sphere is taken into consideration, i.e. $6 \times 6 = 36$ views during sample image acquisition and construction of prototypical view database. This step helps reduce size of database and facilitates recognition process. To compensate for the missing of symmetrical views in database, original testing image and its image inversion are regarded equally as input to recognition.



a b

Figure. 1. Region of landing gear in sample images.

Relative position of a component to plane is fixed. We hereby represent the region enclosing interesting component by a cube. The coordinates of search region in the frontal view is determined by human interaction. For sample images taken from other views, the search region undergoes the same 3D transformation with the aircraft. This helps locate search region in sample image created by the software. Given that the pose of plane is known, the 3D position of search region can be computed with respect to the frontal view, and the 2D projection to an image as a polygon is easily located. The 2D projection, represented by blue polygon in Fig. 1, can be derived numerically from (1) and (2).

Plane model is located in 3D world frame O-XYZ, the X and Y are parallel to the x and y in frontal view (Fig. 1a), and O is plane model center. If the object is transformed by rotation R_m, translation T_m and scale S_m, region point coordinate changes from A to A_{m3d}.

$$A_{m3d} = S_m \cdot (R_m \cdot A + T_m) \tag{1}$$

The 2D projection of search region on training image is

$$A_m = \begin{bmatrix} Ax \\ Ay \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot A_{m3d} \tag{2}$$

B. Aircraft Recognition

Feature-based matching algorithms extract and select image features before matching, which reduces size of image information. A proper feature could maintain attributes of position, rotation and scale of an image. So the feature-based recognition is fast, easily computed and robust. When the transformation between images is unknown, feature-based recognition should be used.

3D recognition method in this paper consists of preprocessing, feature extraction and classifier training. After translation, rotation and scale transformation, fast and precise recognition is difficult. Fourier Descriptor [11] is robust in shape description under the three transformations. In this paper, binary image is derived by preprocessing. Fourier Descriptors are applied to shape feature extraction. For classifier, we use Euclidean distance to find the best match with the shortest distance. The first and third steps will not be discussed in this paper.

The basic idea of Fourier Descriptors is one dimensional Fourier transformation. Object shape, represented by 2D coordinates of its boundary points $\{(x_i, y_i)\}$, is depicted by Fourier coefficients. The analysis of contour in frequency domain is Fourier Descriptor. Because of the reversibility of Fourier transformation, the descriptors at this stage could recover the contour of object with finite numbers of Fourier coefficients. Detail information could be lost more or less according to size of the descriptors.

$$Z_k = DFT(\{x_i, y_i\}) = \frac{1}{N}\sum_{i=0}^{N-1}(x_i + j \cdot y_i) \cdot e^{-j2\pi ik/N}$$
$$k = 0,1,\ldots,N-1 \tag{3}$$

where $\{x_i, y_i\}$ denotes boundary point set, N is the number of contour points.

A descriptor should be translation, scale and rotation invariant to qualify for matching or recognition. Coefficients in (3) must be normalized.

$$d_k = \frac{\|Z_k\|}{\|Z_1\|} \qquad k = 1,\ldots,N-1 \tag{4}$$

According to the property of Fourier series, low frequency coefficients capture most of the energy, whereas high frequency coefficients retain details. The outline of an object could be well represented by a few low frequency coefficients. A coarser or finer description is determined by the number of descriptors. This is an advantage of Fourier Descriptor based on frequency domain to other features based on space domain.

Real scene plane images, as well as other images, contain noise. The segmentation result of plane object would always have a contour that is not smooth. In common feature extraction methods, different filtering techniques are needed to preprocess the result to serve as the input for recognition. However, with Fourier Descriptors, one could ignore noise together with unnecessary details by decreasing the number of coefficients in Fourier series.

C. Search Region Estimation by Major Direction Vector

The best matched plane contour is derived from II *C*. To estimate search region on testing plane object through relative position, one must know the 2D transformation between testing image and training images. The transformation includes scaling, translation and 2D image rotation.

For rotation between plane objects in two images, a direction should be defined that is invariant to the object. Common approach is to treat the main axis of plane as its direction. Symmetry of plane structure is used to determine direction of an aircraft in [16]. When the main axis is parallel to y axis, the average coordinates of object region located on a line parallel to x axis is a constant, $\bar{x}(y) \equiv c$. The minimization of the summation of average coordinates with respect to image rotation angle could align plane objects to the same direction. For example, the symmetry criteria is $\sigma_x(R) = \sqrt{\frac{1}{N}\sum_{x\in\{x_i\}}\left|x-\bar{x}\right|^2}$. The rotation angle of plane object is $\theta_{axis} = \arg\min_{0<\theta<\pi} \sigma_x(R)$.

The direction estimation by symmetry deals with frontal views of plane with an error that is smaller than $\frac{\pi}{128}$. However, as arbitrary views of plane are included in the study, symmetry criteria fails in most cases. A more general method of major direction vector based on approximation ellipse is proposed to find rotation between two similar contours.

Approximate ellipse has same normalized second order central moment with binary object image. Major axis of the ellipse is regarded as principal axis of plane. See Fig. 1. The intersection of the major axis of approximate ellipse and contour with the largest distance is taken as start point, denoted as B_M.

Major direction vector $\overrightarrow{M}$ is proposed to find the 2D transformation between testing image and the best matched training image. From preliminary experiments, head of plane and contour center C lie on principal axis of plane. Distance from head to center is usually larger than distance from tail to center. Major direction vector $\overrightarrow{M}$ is defined as the vector from C to B_M. C and B_M are represented by red and blue triangles in Fig. 1.

The norm of $\overrightarrow{M}$ measures the scale of the object.

Coordinates of C are taken as first order central moment of plane object

$$\bar{x} = \frac{\sum_x\sum_y x\cdot f(x,y)}{\sum_x\sum_y f(x,y)} \tag{5}$$

$$\bar{y} = \frac{\sum_x\sum_y y\cdot f(x,y)}{\sum_x\sum_y f(x,y)} \tag{6}$$

For binary image, $f(x,y)\in\{0,1\}$.

Start point of contour, B_M, satisfies

$$\|B_M - C\| = \max\{\|B_i - C\|\} \tag{7}$$

$$\frac{x_{Bi}.a + y_{Bi}.b + c}{\sqrt{a^2+b^2}} < \varepsilon \tag{8}$$

where $a\cdot x + b\cdot y + c = 0$ is major axis of approximate ellipse, ε is taken as 2 pixels empirically and $\{B_i\}$ is boundary point set satisfying (8).

For the first step of relative position estimation, major direction vector finds the transformation from matched image to testing image by equations (9)-(11). Then the transformation is used to determine search region in testing image as in equation (12).

$$\theta = \theta_t - \theta_m \tag{9}$$

$$R = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{10}$$

$$S = \frac{\|\overrightarrow{M}_t\|}{\|\overrightarrow{M}_m\|} \tag{11}$$

$\overrightarrow{M}_t$ and $\overrightarrow{M}_m$ are major direction vectors of testing image contour and the best matched contour respectively; θ_t and θ_m are corresponding angles of $\overrightarrow{M}_t$ and $\overrightarrow{M}_m$.

$$A_t = S\cdot R\cdot(A_m - C_m) + C_t \tag{12}$$

where A_t is a point in search region in testing image, and A_m is a point in search region in sample image with known coordinates from (2).

III. Experimental Results

A. Noise Immunity of Shape Feature

To test the robustness of Fourier Descriptor, recognition experiment of plane contours with added noise is performed. The noise is random distribution.

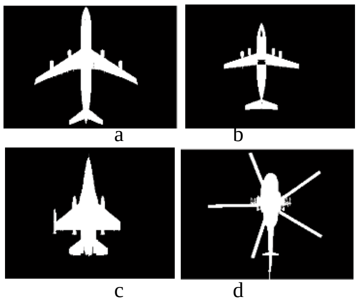


Figure. 2. Four plane models for recognition.

Fig. 3 shows testing image contours after introducing noise, value of noise is 1% to 30% of first central moment of the object, denoted as 1 to 30. Fig. 4 is recognition result of 4 models with different noise. Recognition rates of all models are above 60% with noise index below 29. Contours of model (c) and (d) in Fig. 2 are more distinctive compared to (a) and (b). The recognition rates of (c) and (d) are higher in general, above 80% with noise index below 29.

In Fig. 3, the noise is very heavy with index above 20. The recognition rates of four models are fairly well. In real problem, the testing image regions derived from segmentation and other techniques normally have much smaller noise.

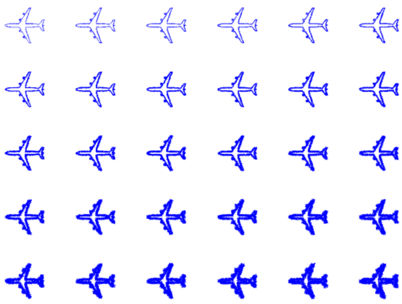


Figure. 3. Introducing noise to contours.

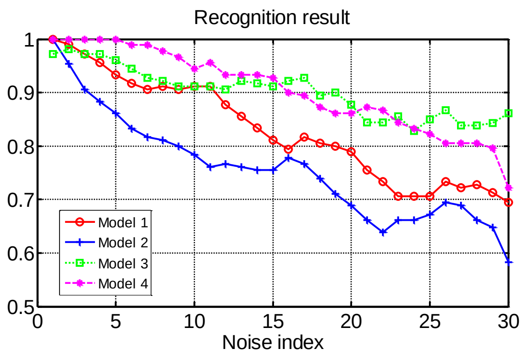


Figure. 4. Recognition rate with different noise.

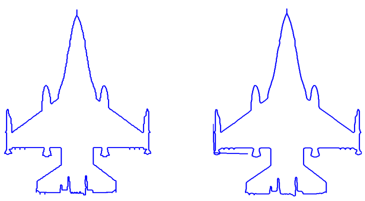


Figure. 5. Original contour and contour after average filtering.

B. Recognition with Smoothed Contour

During this test, real scene testing images are applied to recognition. Filtering is applied to object contours after segmentation or boundary detection. The span of average filter in Fig. 5 is 5. The coordinate of a contour point is the average of itself and 4 nearest points on the contour. The effect of filtering on recognition is shown in Table I.

Recognition rate is the average of rates under 30 noise indexes with each repeating 10 times. The recognition rate of filtered contour means that contours in both the database and testing images are smoothed. Recognition rates of models increase by 8% to 22% after filtering. This means the original contour contains noise and unnecessary details. Although

TABLE I: Recognition Rate Comparison Between Original and Filtered Contours

Type	A340 (a)	BA146 (b)	F16 (c)	Mi17 (d)
Original	0.6171	0.6722	0.8117	0.6878
Filtered	0.8296	0.7565	0.9009	0.9056

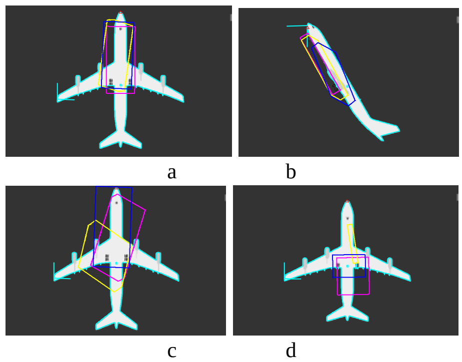


Figure. 6. Search region estimation and detection error.

Fourier Descriptor could reduce noise and detail by adjusting its number of coefficients, filtering is needed to deal with real scene image recognition problem. Experiments in III A and III C are all based on filtered contours.

C. Plane Recognition and Landing Gear Detection

During recognition stage, shape features of objects for detection are matched to all features in database. For relative position estimation stage, the best matched sample after recognition is used to locate search region as in (12). The 2D transformation is determined by major direction vector. Basic image processing operations are then applied to finding landing gear in the region.

Fig. 6 illustrates our recognition and detection result. The red, yellow and blue polygons represent search regions derived from the best matched samples, the runner-ups and the second runner-ups. (a) and (b) are correct recognition results. (c) and (d) show recognition error and detection failure. The recognition and detection rates are computed with the best match in this paper.

In Fig. 7, real scene A380 images are used as testing samples for recognition and component detection research.

(a), (b), (c) are results of correct recognition and detection. (d) is a case of detection error.

Real scene testing images have a scaling of 0.3 to 3 with respect to training images. Since the range is much larger than the scaling of testing images in Table I, the plane recognition rates are generally lower. Because of similarity between different plane models, pose and main axis of object could be matched correctly in some cases even if recognition result is wrong. The search region of testing object overlaps that of the matched one. This accounts for higher detection rate compared to recognition rate in Table II.

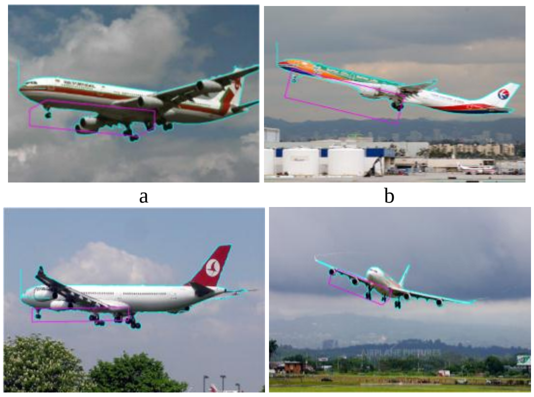


Figure. 7 Real scene recognition and detection result of A380.

TABLE II: PLANE TYPE RECOGNITION AND COMPONENT DETECTION

Type	A340 (a)	BA146 (b)	F16 (c)	Mi17 (d)
Plane recognition rate	0.8056	0.7778	0.6384	0.6944
Gear detection rate	0.8148	0.7870	0.7685	0.8426

IV. CONCLUSION

The paper proposes a detection method by 3D recognition and relative position estimation aiming at inspection of components without distinctive image features. Major direction vector is proposed to locate search region through relative position estimation. Only basic image processing technique, such as thresholding methods or watershed, is needed to find the exact component. Section III *A* shows that model (c) and (d) in Fig. 2 achieve recognition rate higher than 80% under noise index 28. The result of Section III *B* indicates that the average filtering increases recognition rate to 80%-90%, comparable with SfM method in [7], whereas 3D reconstruction in SfM is replaced by a database of 2D views. Final detection result in Table II shows the proposed two-stage algorithm has a detection rate within the range of 75% to 85%.

Due to the complexity of real scene plane images acquisition in large numbers, synthesized images with rotation, translation and scale transformation, as well as noise, are used for testing. Future work will build and test the database with real scene plane images. In addition, image segmentation will be introduced to serve as an input to recognition.

ACKNOWLEDGEMENT

The author would like to thank Dr. Jie Jiang for the beneficial discussion and support of this paper.

REFERENCES

[1] S. Osowski and D. D. Nghia, “Fourier and wavelet descriptors for shape recognition using neural networks—a comparative study,” *Pattern Recognition*, vol. 35, pp. 1949-1957, 2002.

[2] R. Bergevin and J. F. Bernier, “Detection of unexpected multi-part objects from segmented contour maps,” *Pattern Recognition*, vol. 42, pp. 2403-2420, 2009.

[3] K. Arbter, W. E. Snyder, H. Burkhardt, and G. Hirzinge, “Application of affine-invariant fourier descriptors to recognition of 3-D Objects,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, July 1990.

[4] Y. Du, H. Zhang, and Q. Pan, “Three-dimensional aircraft recognition using moments,” *Journal of Data Acquisition & Processing*, vol. 15, September 2000.

[5] H. Wang, Z. Zhuang, and Q. Zhang, “Study on recognition of aircraft status based on dynamic infrared images,” *Infrared and Laser Engineering*, vol. 32, no. 1, February 2003.

[6] I. Gordon and D. G. Lowe, “What and where: 3D object recognition with accurate pose,” in *Toward Category-Level Object Recognition*, Springer Berlin Heidelberg, 2006, pp. 67–82.

[7] E. Hsiao, A. Collet, and M. Hebert, “Making specific features less discriminative to improve point-based 3d object recognition,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2653-2660.

[8] C. M. Cyr and B. B. Kimia, “3D object recognition using shape similarity-based aspect graph,” in *Proc. IEEE International Conference on Computer vision*, vol. 1, 2001, pp. 254-261.

[9] D. G. Lowe, “Distinctive image feature from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, pp. 91-110, 2004.

[10] C. Harris and M. Stephens, “A combined corner and edge detector,” *Alvey vision conference*, pp. 15-50, 1988.

[11] C. T. Zahn and R. Z. Roskies, “Fourier descriptors for plane closed curves,” *IEEE Transactions on Computers*, vol. 100, pp. 269-281, 1972.

[12] T. Wang, W. Liu, J. Sun, and H. Zhang, “Using fourier descriptors to recognize object’s shape,” *Journal of Computer Research and Development*, vol. 39, December 2012.

[13] S. Savarese and F. Li, “3D generic object categorization, localization and pose estimation”, *ICCV*, 2007, pp. 1 - 8

[14] X. F. Chen, “Research on object representation and recognition techniques based on skeleton,” PhD Dissertation, School of National University of Defense Technology, October 2004.

[15] Y. J. Du, H. C. Zhang, and Q. Pan, “Three-dimensional aircraft recognition using moments,” *Journal of Data Acquisition and Processing*, vol. 15, no. 3, September 2000.

[16] Z. M. Qian, “Researches on image segmentation approaches,” Master Dissertation, School of National University of Defense Technology, November 2010.