

AERIAL OBJECT RECOGNITION ALGORITHM BASED ON CONTOUR DESCRIPTOR

V.V. Strotov^{a,*}, P.V. Babayan^a, S.A. Smirnov^a

^a Dept. of Automation and Information Technologies in Control, Ryazan State Radio Engineering University,
390005, Gagarina Str. 59/1 Ryazan, Russia (strotov.v.v, babayan.p.v, smirnov.s.a) @rsreu.ru

Commission II, WG II/5

KEY WORDS: vision system, object recognition, Euler angles, geosphere, contour descriptor

ABSTRACT:

This paper describes the image recognition algorithm for on-board and stationary vision system. Suggested algorithm is intended to recognize the aerial object of a specific kind using the set of the reference objects defined by 3D models. The proposed algorithm based on the outer contour descriptor building. The algorithm consists of two stages: learning and recognition. Learning stage is devoted to the exploring of reference objects. Using 3D models we can gather set of training images by rendering the 3D model from viewpoints evenly distributed on a sphere. Sphere points distribution is made by the geosphere principle. Gathered training image set is used for calculating descriptors, which will be used in the recognition stage of the algorithm. The recognition stage is focusing on estimating the similarity of the captured object and the reference objects by matching an observed image descriptor and the reference object descriptors. The experimental research was performed using a set of the models of the aircraft of the different types. The proposed orientation estimation algorithm showed good accuracy in all case studies.

1. INTRODUCTION

The aerial object recognition is one of most interesting and complicated tasks that can be solved using image processing (Duda, 1976) and (Shapiro, 2001). In this work we understand the aerial object recognition as a determination of its type. The set of aerial object types includes various kinds of airplanes, helicopters and UAVs.

The object recognition approaches include model-based approaches, shape-based approaches, and appearance-based models (Roth, 2008). Model-based approaches try to represent (approximate) the object as a collection of three dimensional, geometrical primitives (boxes, spheres, cones, cylinders, generalized cylinders, surface of revolution) whereas shape-based methods represent an object by its shape/contour. In contrast, for appearance-based models only the appearance is used, which is usually captured by different two-dimensional views of the object-of-interest. Based on the applied features these methods can be sub-divided into two main classes, i.e., local and global approaches.

The goal of this work is developing of algorithms for onboard vision systems. Here are the main requirements for proposing algorithms:

- Possibility of real-time implementation with limited computational resources;
- Target object size, position and pose invariance;
- Workability in wide range of object capturing distances.
- Object set invariance (with known object 3d-models);

The developed algorithm can be applied for different task solving:

- Air traffic control;
- Aircraft collision avoidance;
- UAV navigation and control.

The most recent approaches for the aerial object recognition task solving are based on object contour classification (Kumar, 2005) and (Belongie, S.. 2002). They can be divided on three groups.

1. Fourier descriptors based methods (Wong, 2007) and (Bonciu, 1998).
2. Object contour selection methods based on Hu invariant moments (Hu, 1962).
3. Methods based on outer object contour comparison (Alpatov, 2016).

The listed methods are invariant to image shift and rotation. However, the Fourier descriptors based methods and methods based on Hu invariant moments show lower precision and require more computational resources in comparison with outer object contour based method. The last method is insensitive to additive noise, but sensitive to object illumination variation.

Thereby, the Methods based on outer object contour comparison are the most suitable for object contour information extraction. The goal of this work is developing the object recognition algorithm based on the object contour information.

* Corresponding author

2. THE RECOGNITION ALGORITHM

2.1 3D pose representation

In this work, we choose Euler angles (α, β, γ) for 3D pose representation (Bekir, 2007). There are various conventions of Euler angles (axes rotation sequences).

We choose XYZ convention. This convention means successive rotations around the axes X, Y, Z in the global coordinate frame. Global coordinate frame is the frame attached to the observer (camera). Consequently, as the Z axis in global frame coincides with the camera optical axis, than the third Euler angle γ corresponds to the camera rotation around the optical axis.

2.2 The general approach to pose estimation problem

The proposed algorithm consists of two stages. The *first stage* (learning stage) executes once when the 3D models of an reference object set is given. It has high computational complexity and usually should be executed on a high-performance computer. The learning stage consists of:

- 1) Sphere discretization – obtaining uniformly distributed set of sphere points (set of discrete values of the first two values of the Euler angles);
- 2) Training images rendering;
- 3) Descriptor calculation for every training image.

The *second stage* (recognition stage) is executed by the request. It requires significantly less computational effort than the first stage. The similarity values for every reference object are calculated by matching descriptor of observed image with the descriptors of training images from the set. The block diagram of proposed recognition approach (both stages) is illustrated in figure 1.

Taking this approach as a basis, we can build different recognition algorithms, taking different types of image descriptors and different ways of descriptor matching. In the next section we describe sphere discretization process.

2.3 Sphere discretization

The proposed algorithm requires training images rendering. The set of training images should be full enough to cover all possible object poses. At the same time the set should not be redundant to minimize calculations.

To get such a training set we put object of interest in the center of imaginary sphere. Putting camera in different points on the sphere surface we can get object images from these sphere points.

These points should be uniformly spread on a sphere so we have to solve sphere discretization problem (Saff, 1997). Every sphere point determines two first Euler angles α and β , keeping spare the third angle γ .

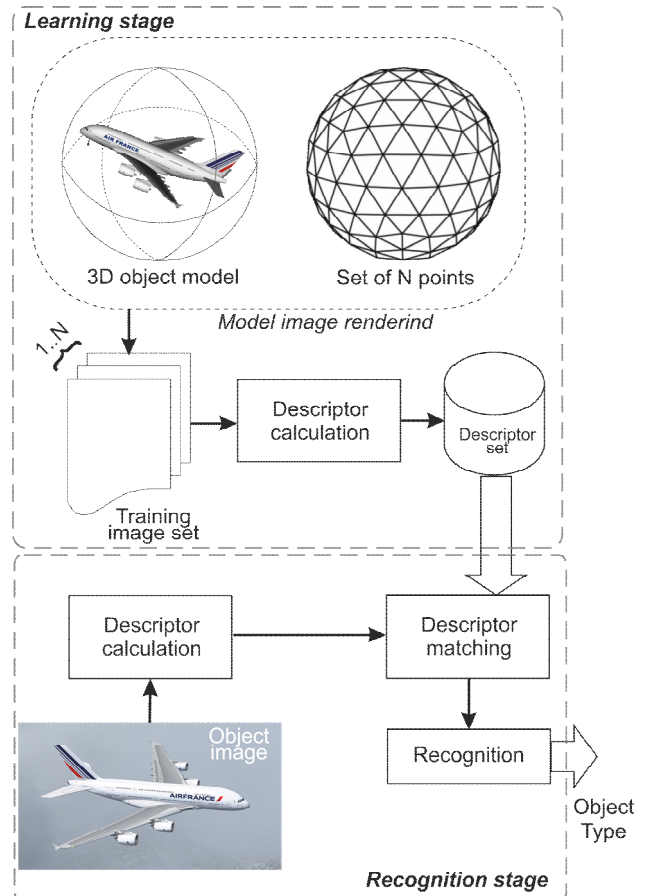


Figure 1 – General block diagram of the proposed algorithm

We consider three approaches for the point distribution.

- 1) Distribution of points in the intersections of parallels and meridians of a sphere. This solution is easy enough to implement, but it lacks the distribution uniformity. Denseness of points is high at the poles of the sphere, and low – at the sphere equator.
- 2) Spiral curve point distribution.
- 3) The geosphere principle. The particular case of geosphere is icosahedron. Points of a geosphere are calculated by division of the icosahedron into 4^{n-1} equilateral triangles, where n stands for factor of geosphere (figure 2).

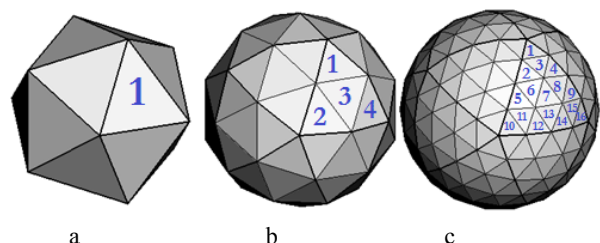


Figure 2. Geosphere building.

- (a) Geosphere factor - 1, number of the vertexes - 12
- (b) Geosphere factor - 2, number of the vertexes - 42
- (c) Geosphere factor - 3, number of the vertexes - 92

An experimental research was performed to compare sphere discretization methods. The results of the research are provided at Figure 3. Figure 3,a-c shows discrete spheres, calculated by different discretization methods and corresponding histograms of angular distances between sphere points (in degrees).

Using obtained histograms we can make a conclusion that distribution based on the geosphere principle provides the most uniform distribution of points on a sphere. It should be noticed that the histogram and variation of the angular distances between the geosphere points differs from theoretical data because of the errors occurring in recursive algorithm of geosphere point's calculation.

2.4 Descriptor matching

The algorithm proposed is based on description of extracted image contour. After binary thresholding the query image, we can extract external image contour. Points of the contour are translated into polar coordinate frame with the frame center in the object centroid. Obtained vector of polar coordinates is discretized and subjected to the median filter. Resulting vector is the contour descriptor. Steps of calculation of contour descriptor are illustrated in the figure 4.

The most probable pose estimation is calculated as a result of matching the contour descriptor of query image with training descriptors. Descriptor matching is performed by the criterion function:

$$f_{crit}(j) = \min_{s \in S} \left[\sum_{i=1}^{N_D} (D_0(i-s) - D_j(i))^2 \right], \quad (1)$$

$$S = \{s : s = \overline{1, N_D}\}$$

where D_0 is the query image descriptor, D_j is the descriptor of current training image, N_D is the descriptor's size, i is an index of point of descriptor, j is an index of training descriptor, and s is the value of circular shift of descriptor. This criterion function provides rotation invariance of the descriptor. Index j of training descriptor corresponds to a geosphere point, therefore it determines Euler angles α and β . Let s_0 to be the shift value s that gives the minimum value to the expression in square brackets in (1):

$$s_0(j) = \arg \min_{s \in S} \left[\sum_{i=1}^{N_D} (D_0(i-s) - D_j(i))^2 \right], \quad (2)$$

$$S = \{s : s = \overline{1, N_D}\}$$

Hence the value of angle γ is calculate by the formula

$$\gamma = \frac{2\pi s_0}{N_D}.$$

(3)

As a result of calculating criterion function (1) for every training descriptor we get vector of values of criterion function

$$M = (f_{crit}(j) | j = \overline{1, N_g}) \quad (4)$$

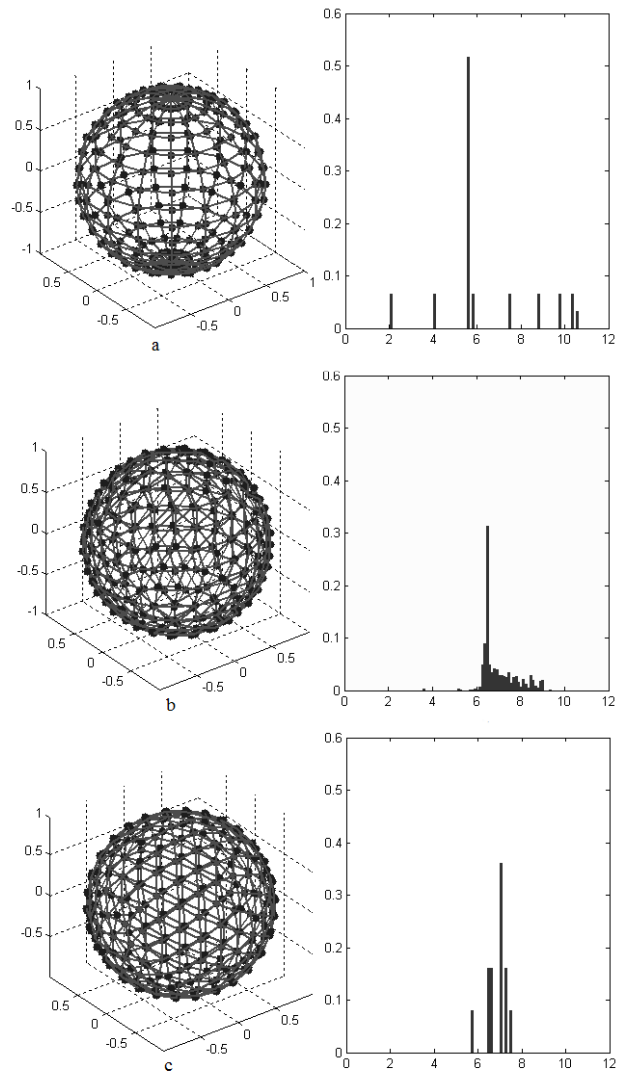


Figure 3. Sphere discretization.
Left figures – sphere examples, right figure – histograms of the angles between the neighbor points
(a) Parallels and meridians
(b) Spiral distribution
(c) Geosphere distribution

The measure of the similarity between captured object and the k -th reference object is value R_k . It can be defined as the minimal distance between sets of the object descriptors as:

$$R_k = \min_j (f_{crit}(j)), \quad j = \overline{1, N_g}. \quad (5)$$

The recognition is processed by finding the least value of the R_k :

$$match = \arg \min_k (R_k), \quad k = \overline{1, K}, \quad (6)$$

where $match$ – the index of the most similar reference object, K – the cardinality of the reference object set.

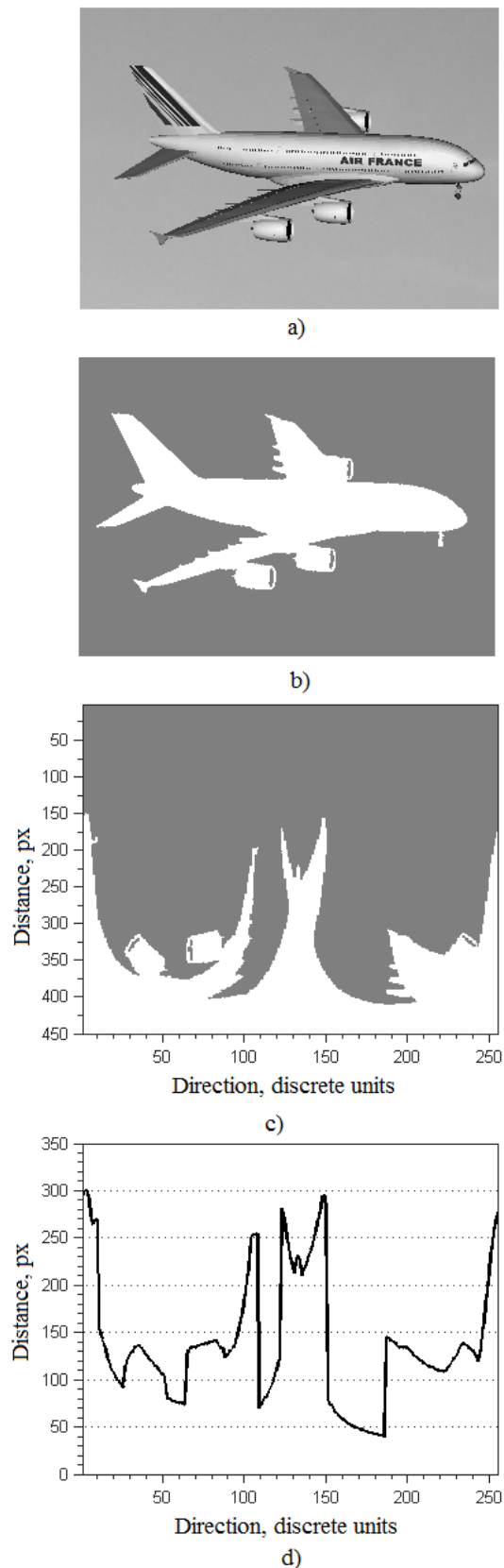


Figure 4 – Steps of contour descriptor calculation:
 a – the input query image of aircraft,
 b – binary image, c – binary image translated in polar coordinate frame, d – external contour descriptor

As the suggested algorithm is based on image contour description, we meet the ambiguity problem. It happens when calculated descriptor corresponds to more than one orientation. For instance, topside and underside views of an airplane will provide equal object contours and hence descriptors. This problem must be taken into account in case of solving orientation estimation task, but it is not important in case of recognition task solving.

3. EXPERIMENTAL RESEARCH

The goal of the experimental research was to study the performance and accuracy of the proposed algorithm. The experiments were carried out on 10 natural image sequences. The sequences were captured in visual and IR bands. The lengths of the sequences were about 20 seconds at frame rate 25 Hz. The minimal areal object area is 500 pixels. The maximal areal object area is less than 15% of the image area. The objects types were airplanes, helicopters and UAVs.

The object recognition algorithm was aggregated with aerial object detection and parameter estimation algorithm. The brightness of the points fits to the range from 0 to 1. Image sensor was fixed but they were imperceptibly moving. The examples of the observed frames are shown on Figure 5.

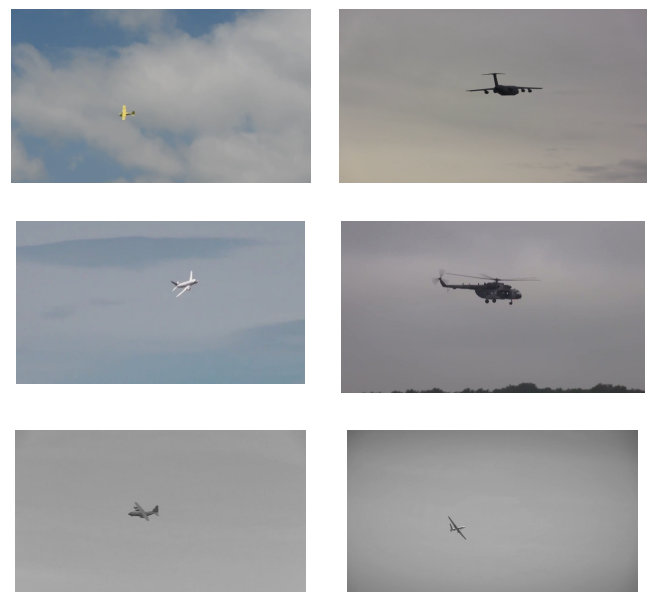


Figure 5. The sample frames of the test video sequences

The reference object base includes 17 objects. The objects were defined by the 3D models. The sets of the reference images were rendered for every model. The factor 3 geosphere point distribution was used (92 points). The light source was situated in front of the object. The samples of the object binary images of the different types are shown of figure 6.

The quality of the object recognition was estimated using the true positive recognition ratio. This value was averaged on the entry test video set. The results of the experiments are shown in the Table 1.

№	Video length, frames	Object	Estimated object class	True positive ratio, %
1	104	Airbus 380	«Airplane»	96,1
2	226	LockheedC-130	«Airplane»	92,9
3	243	Mi-172	«Helicopter»	98,3
4	169	MQ9	«UAV»	86,9
5	104	MQ1	«UAV»	80,8
6	357	Mi-8	«Helicopter»	86,2
7	139	IL 76	«Airplane»	100
8	375	Lockheed C-5	«Airplane»	100
9	318	Cessna 172	«Airplane»	95,6
10	355	Sukhoy Supejet 100	«Airplane»	100
Average				93,7

Table 1. Results of the experimental examination

4. CONCLUSION

A contour-based approach to solve aerial object recognition problem was considered in this paper. The approach is based on outer object contour building. The algorithm consists of two stages: learning and recognition. The learning is performed on training images rendered from 3D models of reference objects in different poses uniformly spread on a sphere. On the recognition the most similar reference object is searched using the outer contour descriptor comparison. The proposed algorithm shows the high accuracy, noise resistance and low computational complexity. The research can be used to develop and enhance of real-time image processing and analysis systems.

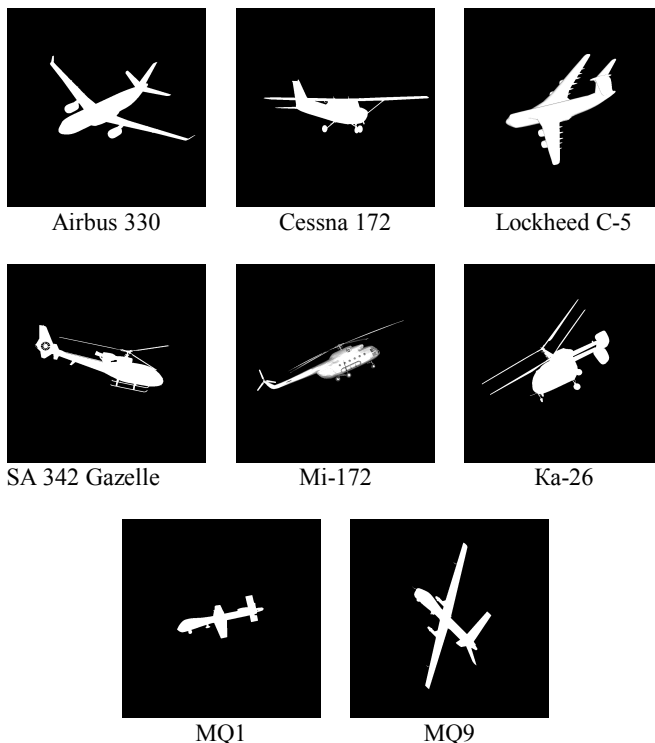


Figure 5. The sample binary images of the reference objects:
Top row: objects of class «Airplane»,
middle row: objects of class «Helicopter»,
bottom row: objects of class «UAV».

ACKNOWLEDGEMENTS

This publication has been prepared as part of research carried out by the Ryazan State Radio Engineering University in accordance with the state contract 2.7064.2017/BCh.

REFERENCES

- Alpatov B.A., 2016. “Contour-based object orientation estimation” in Proc. of SPIE Real-Time Image and Video Processing, Vol. 9897.
- Belongie, S.. 2002. “Shape matching and object recognition using shape contexts” in IEEE Transactions on Pattern Analysis and Machine Intelligence – V. 24. – №. 4. – Pp. 509-522.
- Bekir, E. 2007. *Introduction to Modern Navigation Systems*, World Scientific Publishing Co. Pte. Ltd.
- Bonciu, C., et al. 1998. A Fourier-Shannon approach to closed contours modelling. *Bioimaging*. – Vol. 3(6). – C. 111-125
- Duda, R., et al.. 1976. *Pattern recognition and the analysis of scenes*. Mir, Moscow.
- Hu, M. K., 1962. Visual pattern recognition by moment invariants. *Information Theory*, IRE Transactions on. V. 8. – №. 2. – Pp. 179-187.
- Kumar, B.V.K.V. et al. *Correlation pattern recognition*. Cambridge University Press, 2005.
- Roth, P. M. et al. 2008. Survey of appearance-based methods for object recognition. Inst. for Computer Graphics and Vision, Graz University of Technology, Austria, Technical Report ICGTR0108 (ICG-TR-01/08).
- Saff E. et al. 1997. Distributing many points on a sphere. *The Mathematical Intelligencer*, Vol. 19, No. 1, 5-11
- Shapiro, L., et al. 2001. *Computer vision*. Prentice Hall.
- Wong W. T., et al. 2007 Shape-based image retrieval using support vector machines, Fourier descriptors and self-organizing maps //Information Sciences. №. 8. – P. 1878-1891.