Air object recognition by the normalized contour descriptors

Valentyn Yesilevskyi, Andrew Tevyashev, Anton Koliadin

National University of Radio Electronics, Nauky ave.,14, Kharkiv, 61166, Ukraine

Abstract

This paper consider research into the methods for recognizing the type of an air object on a digital image acquired from video monitoring system. A method has been proposed that is applied on a feature vector built on the basis of a Fourier transform for the sequence of coordinates of its two-dimensional contour. This makes it easier to solve the classification problem owing to a more compact arrangement of the multidimensional feature vectors for similar air objects. The architecture of an air situation video monitoring system has been suggested, which includes an image preprocessing module and a module of neural network. Preprocessing makes it possible to identify an object's contour and build a sequence of normalized descriptors, which are partially independent of the spatial position of the object and the contour processing technique. Proposal method is easy in realization and do not require significant computational resources due take into consideration the specificity of recognizing objects in 3 dimensional. This research has shown that the reported results make it easier to train a neural network and reduce the hardware requirements for solve the task of air situation video monitoring.

Keywords

Air object recognition, contour analysis, Fourier descriptors, neural network

1. Introduction

Solving tasks of detection and recognition of air objects fast at high quality is of great importance for civilian applications in air traffic control, airport air situation monitoring, as well as military activities. Currently, the list of types of aircraft objects has been significantly expanded through the use of unmanned aerial vehicles (UAVs), quadcopters, cruise missiles, and helicopters. Optical video surveillance systems are increasingly being used to detect aerial objects. Video monitoring reduce the dimensions of detection systems and as well avoids problems related to the masking of characteristics of air objects in the radar detection area (stealth technology, small size, etc.). Video technologies require to solve tasks into the field of digital video processing and automated object recognition.

Solving them is typically based on the use of deep convolutional networks. Such a solution is too universal and not taking into consideration that in the area of recognition of the type of air objects, the contours of the object yield enough information to solve the problem. But the issue is that the resulting numerical characteristics of the contour should be invariant relative to the geometric distortion (displacement, orientation, scale) of the object's image. For air objects, this is especially relevant due to the three degrees of freedom in determining the position and is an important area to research.

One of the most effective and commonly used approaches to numerical description of the geometric shape of a flat object's contour is the application of a Fourier transform procedure. This approach is particularly interesting because it generates a unique one-dimensional identification sequence of the standard size, termed as Fourier descriptors, for all examined objects.

The deep convolutional networks that are used in laboratory image recognition experiments are too heavy both at the training stage and at the operational stage. For actual application, including mobile devices, methods that require fewer computing resources are in demand. This is confirmed by works [1–2] that address the use of object recognition methods based on feature

EMAIL: valentyn.yesilevskyi@nure.ua (A. 1); andrew.teviashev@nure.ua (A. 2); anton.koliadin@nure.ua (A. 3)
ORCID: 0000-0002-5935-1505 (A. 1); 0000-0001-5261-9874 (A. 2); 0000-0001-5552-5080 (A. 3)

vectors, including Fourier descriptors, acquired from a contour analysis. However, known solutions involving these methods either do not use information about phase from vectors complex vectors or do not parse the contours of objects into classes corresponding to three different projections. This suggests the relevance of our research into recognition air objects in an arbitrary spatial position, based on the Fourier descriptors of the contour.

2. Problem statement

Study [3] gives a detailed overview of the significant advances in the use of deep convolutional neural networks to recognize arbitrary images. However, as shown in papers [4, 5], deep learning methods require a large number of training sets and significant computer time to train a network.

It is shown in [6] that the use of transfer learning may become an option to overcome the heaviness of deep network training from scratch. This method assumes that a pretrained universal network additionally learns from specific images of air objects. The application of transfer learning makes it possible to bypass the limitations associated with the training time. However, some questions remain open. First, there is the issue of the redundancy of a universal approach based on deep neural networks when it comes to the recognition of types of air objects. Second, there is an issue related to the division of object classes, each of which is represented by too different (due to spatial orientation) images of objects on a flat image.

It should also be taken into consideration that explaining the recognition process in deep networks is based on highlighting the hierarchy of attributes in the image. This means that the network learning process involves excessive training the network, including for generating the contours' attributes, although this can be solved more effectively by classical methods of image processing.

Describing an image of an object by its contour is sufficient for the task of recognizing the types of air objects; it utilizes much less information than when analyzed using deep neural networks, which allows for a series of advantages. Studies [7, 8] report various methods for the mathematical notation of a contour; work [9] shows the use of these methods to recognize the types of air objects.

When recognizing 3-dimensional objects by their 2-dimensional image, there is an issue related to deriving a numerical descriptor invariant relative to the orientation of the object and its size in the image. The most interesting for description the contour of an air object is the use of Fourier descriptors. This method based on a discrete Fourier transform to the function that describes the contour of the object. As shown in papers [7, 8], this method is based on the representation of a closed flat curve by a sequence of points whose coordinates are considered complex numbers. Similar to how the spectrum of an audio signal, derived from a its Fourier transform, identifies this signal, the Fourier descriptors identify the closed contour of the flat shape.

Fourier descriptors partially resolve the issue of invariance in relation to the rescaling and rotation in the image plane. However, direct use of all the information derived from a Fourier transform is not possible as the phase component of the descriptors depends on the choice of the starting point of the contour. This issue requires separate research.

The task of recognizing aircraft types based on feature vectors derived from the Fourier descriptors has been solved since the 1980s by using a variety of classification methods. Work [10] applies methods of correlational analysis, distance assessment, support vector machine; paper [11] explored the use of neural networks to solve the task of categorizing objects by their descriptors of a Fourier contour.

A review of the results of the above studies reveals the following shortcomings in existing approaches to solving a task of air object recognition. Deep convolutional network methods are excessively heavy and general-purpose. The methods of network training do not take into consideration the possibility to simplify learning when dividing the training sample into subclasses corresponding to three different projections.

All this suggests that it is appropriate to construct an air object recognition method that would combine the advantages of a contour analysis based on Fourier descriptors and the specific training method of neural networks in order to overcome the above shortcomings

3. Studying the dependence of changes in the descriptors of the contour of air objects on the angle of rotation

This study is assumed that the image of an air object can be represented as an ordered sequence of $z(k) = (x_k, y_k)$, $k = 0,\ldots, N–1$ points that describe the contour.

The application of a Fourier discrete transform to this sequence generates a unique one-dimensional identification sequence of standard size values called Fourier descriptors, which possesses a series of interesting properties that are studied in detail below.

3.1. Exploring the properties of Fourier descriptors

A one-dimensional Fourier transform of the function $f(t)$:

$$F(\omega)=\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{1}$$

makes it possible to derive a continuous spectrum of this function in the frequency domain normally used to analyze time signals. For the case of the discrete sequence N of points $z(k)$ of the examined signal (1) takes the form of a discrete Fourier transform and leads to the calculation of the discrete spectrum:

$$F(n)=\sum_{k}^{N-1} z(k)e^{-i\frac{2\pi nk}{N}}, \tag{2}$$

where $n=-N,\ldots,0,1,\ldots,N-1$.

It is important that this transform allows for a reverse operation – an inverse discrete Fourier transform:

$$z(k)=\frac{1}{N}\sum_{n}^{N-1} F(n)e^{i\frac{2\pi nk}{N}}, \tag{3}$$

that restores the original sequence.

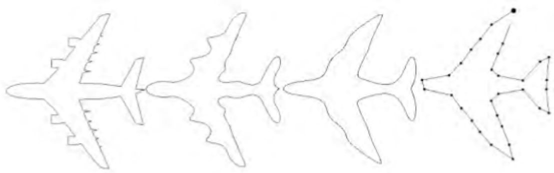
The sequence of Fourier descriptors possesses a series of important properties that explain their use in describing the contours of objects in two-dimensional images: invariant under shift, rotation, and scaling, as well as orderliness.

However, Fourier descriptors are sensitive to the choice of the starting point in Fourier transform processing. This effect makes it difficult to directly use the phase component of the descriptors to identify the contour.

There are also problems with arbitrary affine transformation. Such a transformation is one of the simplest, which roughly describes the distortion of the shape of a flat image of a contour when a 3-dimensional object is at an arbitrary angle relative to the plane of the camera.

The one of important useful feature in terms of contour identification is the orderliness of the descriptors by the degree of their importance to the image. This makes it possible to discard (zero) the high-frequency components without losing visual recognizability. In the sequence of the Fourier descriptors for a flat shape, the information about the shape of the contour is delivered by the first few elements of the sequence.

An inverse Fourier transform makes it possible to assess the degree of recognition of the shape at zeroed high-frequency components. The image contour shown in Figure 1, *a* contains about 2,000 points. Accordingly, a Fourier discrete transform would result in the same number of descriptors. The inverse transform would result in an accurate contour restoration.



a *b* *c* *d*

Figure 1: Illustration of the image contour filtering effect involving a Fourier transform

The zeroing of the descriptors above some of the filter's chosen boundary frequency results in a decrease in the details in the image at inverse transform while the object's recognition is retained. Figure 1, *b, c* shows the contour restored after zeroing all but 64 and 32 low-frequency descriptors, respectively. Figure 1, *d* shows an image obtained for the same original contour by discarding all descriptors except 32 for the rapid derivation of only 32 points of the restored image with linear interpolation between them.

3.2. Choosing a feature vector

In a general case, the Fourier descriptors $F(n)$ can accept different values for different images of the flat contour even of the same object. Their magnitudes depend on the scale r, the rotation angle φ, and the choice of the contour starting point k_0, as considered above.

Assuming that some reference value of the descriptors $F^*(n)$ has been selected, the calculated $F(n)$ sequence) for the recognized object can be represented as follows:

$$F(n) = re^{i\varphi}e^{-i\frac{2\pi nk_0}{N}} \cdot F^*(n). \qquad (4)$$

One can convert the Fourier descriptors to a form that lacks the influence of these factors. Consider the normalized descriptors according to [12]:

$$N(n) = \frac{F(1+n)F(1-n)}{F^2(1)}. \qquad (5)$$

It is then possible to show that:

$$N(n) = \frac{F^*(1+n)re^{i\varphi}e^{-i\frac{2\pi(1+n)k_0}{N}}F^*(1-n)re^{i\varphi}e^{-i\frac{2\pi(1-n)k_0}{N}}}{\left(re^{i\varphi}e^{-i\frac{2\pi k_0}{N}}F^*(1)\right)^2} = \qquad (6)$$

$$= \frac{F^*(1+n)F^*(1-n)}{\left(F^*(1)\right)^2} = N^*(n),$$

where $N^*(n)$ are the normalized coefficients, corresponding to the reference set of the Fourier descriptors $F^*(n)$.

It is easy to see that the normalized descriptors $N(n)$ (5), unlike Fourier descriptors, are independent of the above factors and can serve as an adequate sequence to describe the reference image of the contour.

3.3. Computational experiment

We have investigated the dependence of Fourier descriptors (2) and the normalized descriptors (5) on the angles of spatial position for four different types of aerial objects: aircraft, unmanned aerial vehicle, helicopter, quadcopter. In accordance with the filtering properties of a Fourier transform shown in Figure 1, it was determined that the first 32 descriptors would suffice to represent the shape of the air objects being examined to solve a classification problem.

One can see (Figure 2) that even similar profiles of the aircraft and helicopter have different «patterns» in terms of the amplitude component for their Fourier descriptors. The phase component, as discussed above, is not representative because of the uncertainty of choosing the starting point of the contour. However, this difference is especially noticeable for normalized descriptors where both the amplitude and the phase component are important.

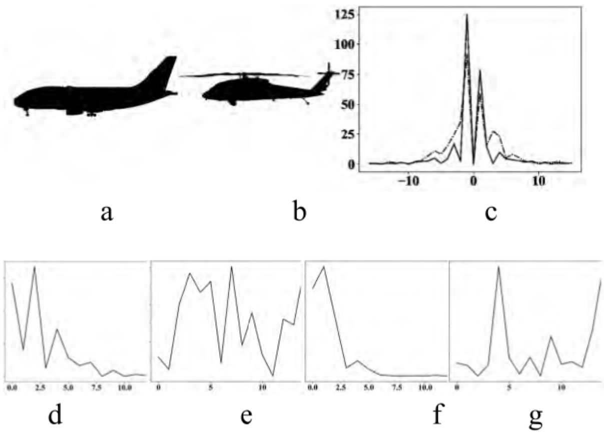


Figure 2. Difference in the shape of the objects represented by their Fourier descriptors and normalized descriptors: a – a binary image of the A380 aircraft profile; b – the same for the Apache helicopter; c – amplitude characteristics of Fourier descriptors (solid line – for A380, dotted line – for Apache); d, e – the amplitude and phase characteristics of the normalized descriptors for A380; f, g – the same for the normalized descriptors for Apache

When recognizing the type of any three-dimensional object in a two-dimensional image, there is an issue of the mutual location of the object and the video camera. Regardless of the classification method the proximity of the view of the contour of the real object and some pattern remains important. This is a particularly important issue for aerial objects that have the freedom to rotate around any of the 3 axes in threedimensional space.

There are a series of simplifications. Specific terms are typically used for aircraft rotation angles (roll, yaw, pitch). One can assume that the video plane is parallel to the XY plane of the system of coordinates of the aircraft Figure 3). At the same time, the aircraft can be simplified to present as a model in the form of flat surfaces (Figure 3, *a*). Then it is obvious that the roll of the aircraft would result in an easy computed change in the contour of the wing projection in the image.

In this case, for each point of this contour relative to the contour at a zero roll, the y-th coordinate would change proportionally to the cosine of the roll angle.

Figure 3 Change in the aircraft projection when its position changes

One can more accurately describe the change in the coordinates of a rotating object. Since of interest is a change in the x and y coordinates, one can record:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\beta & 0 \\ -\sin\alpha\cos\beta & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \qquad (7)$$

where x' and y' are the coordinates of the new position of the arbitrary point of the object at coordinates x and y.

It is easy to show that due to the linearity of a Fourier transform, the proposed normalized descriptors (5) do not depend not only on such a change of angles but, in general, on an arbitrary affine transformation.

Studying the dependence of Fourier descriptors (2) and normalized descriptors (5) on the spatial angles for the examined air objects has shown that when the angle of the roll increases from 0 to 90 degrees, the side image smoothly transitions into the image from below (Figure 4), while the Fourier descriptors and the normalized descriptors are smoothly transformed into the horizontal projection descriptors.

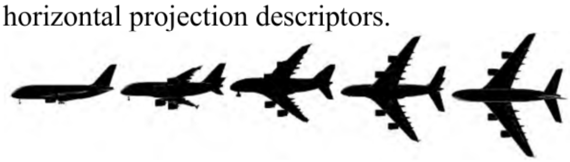


Figure 4 Change in the shape of the A380 aircraft's projection when the angle of the roll changes from 0° to 90°

However, in practice, the representation of an aircraft in the form of flat surfaces is too rough; the descriptors of different frequencies perform differently during this transformation. Figure 5 gives the results of a change in the amplitude of the first few normalized descriptors. Similar studies have been conducted for other types of aircraft (helicopter, UAV, and quadcopter).

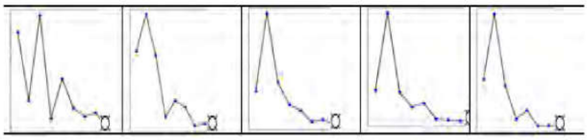


Figure 5 Change in the amplitude of the normalized contour descriptors when the angle of the A380 aircraft roll changes from 0° to 90°

It is generally accepted that the sequence of Fourier descriptors is invariant not only to the change in scale and displacement and only the roll and pitch influence the change in the two-dimensional projection of the three-dimensional silhouette of an aircraft.

Our analysis revealed that in the case when the shooting point is not directly under the plane, even yaw is not a pure image rotation, which can be compensated for when processing the descriptors. Thus, even in the trivial case, if one studies the contour of a cube or parallelepiped, at a yaw rotation the quadrangle contour can change to hexagonal.

The result of studying the contours of 4 types of objects: passenger plane, helicopter, unmanned aerial vehicle, quadcopter has established that the normalized descriptors $N(n)$ (5) even for low frequencies are sensitive to changes in the position of the object in space. That necessitates the use of separately normalized descriptors for each of the three orthogonal projections of the object as class references.

Therefore, three orthogonal projections as separate classes are used when solving a classification task and building a training set in this study, although this increases the number of classes to 11 for the 4 selected types of objects (the quadcopter has two projections that match).

4. Discussion of results of studying

Our results of the convergence of the training procedure indicate that the proposed approach to the selection of normalized contour descriptors as the vectors of object attributes in order to solve the task of recognizing the types of air objects using a neural network makes it possible to build an effective recognition subsystem for the complex of moving air object detection

The application of the proposed method is associated with a series of limitations in terms of its practical use. For example, solving the task implies, first, that an air object is large enough to detect a contour that can be used to calculate 15 normalized descriptors. Second, it is assumed that the image of the object is not distorted by fog or partially closed by clouds. The removal of these restrictions could be partially implemented by the pre-processing of the image and is the subject of further research.

5. Conclusions

A computational experiment involving model images was performed to investigate the dependence of change in Fourier descriptors and the normalized contour descriptors of images of 4 types of air objects on the angle of the object's rotation relative to 3 axes. It is shown that the use of normalized descriptors has a series of advantages over the Fourier descriptors. It has been established that to simplify the recognition task, one needs to parse the training set for each type of object into 3 classes corresponding to 3 orthogonal projections. This makes it easier to solve the classification problem owing to a more compact arrangement of multidimensional feature vectors for shapes with similar images.

Algorithmic maintenance for an air situation video monitoring system has been developed. The task of recognizing the type of an air object is tackled by a pre-trained neural network. The validation detection accuracy is as high as 99 %. This confirms that the proposed method of building a system for recognizing the types of air objects could simplify the requirements for the implementation of hardware while improving the accuracy when solving a task of air situation recognition.

6. References

[1] Strotov V., Babyan P., Smirnov S. (2017). Aerial object recognition algorithm based on contour descriptor. ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. DOI: 10.5194/isprs-archives-XLII-2-W4-91-2017.

[2] Luciano da F. Costa, Roberto Marcondes Cesar Jr. (2009). Shape Classification and Analysis: Theory and Practice, Second Edition. 10.1201/9781315222325.

[3] Neha Sharma, Vibhor Jain, Anju Mishra, An Analysis Of Convolutional Neural Networks For Image Classification, Procedia Computer Science, Volume 132, 2018, Pages 377 – 384. DOI: 10.1016/j.procs.2018.05.198

[4] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. Commun. ACM 2017, 60, 84–90. DOI: 10.1145/3065386.

[5]. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9. DOI: 10.1109/CVPR.2015.7298594, 10.1109/NAECON.2016.7856803.

[6] Valentyn Yesilevskyi, Andrew Teviashev, Anton Koliadin. Transfer learning in aircraft classification. IST-2019. the 8-th International Scientific and Technical Conference September 9 – 14, 2019 Kobleve-Kharkiv, Ukraine p.132 – 135. https://openarchive.nure.ua/bitstream/document/11942/1/3_IST.pdf.

[7]. Chaki Jyotismita, Dey Nilanjan. (2019). A Beginner's Guide to Image Shape Feature Extraction Techniques. 10.1201/9780429287794.

[8] Mingqiang Yang, Kidiyo Kpalma, Joseph Ronsin. Shape-based invariant features extraction for object recognition, Springer, 60 p. 2012. hal-00652422 DOI: 10.1007/978-3-642-24693-7_9.

[9] Hai-Jun Rong, Ya-Xin Jia, Guang-She Zhao, Aircraft recognition using modular extreme learning machine, Neurocomputing, Volume 128, 2014, Pages 166 – 174. ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2012.12.064.

[10] Makarov M.A., Berestneva O.G., Andreev S. Yu. (2014). Solving the problem of describing and classifying the contours of moving objects on video. Bulletin of the Tomsk Polytechnic University. Engineering of Georesources, 325 (5), 77 – 83.

[11] Nguyen, T. T. (2010). Algorithmic and software for shape recognition using Fourier descriptors and a neural network. Bulletin of the Tomsk Polytechnic University. Georesource Engineering, 317 (5).

[12] Granlund, G.H. (1972). Fourier Preprocessing for Hand Print Character Recognition. Computers, IEEE Transactions on. 21. 195 – 201. 10.1109/ TC.1972.5008926.