Using Shape Descriptors for UAV Detection

Eren Unlu, Emmanuel Zenou, Nicolas Rivière

Using Shape Descriptors for UAV Detection

Eren Unlu; ISAE-SUPAERO; Toulouse; France
Emmanuel Zenou; ISAE-SUPAERO; Toulouse; France
Nicolas Riviere; ONERA; Toulouse; France

Abstract

The rapid development of Unmanned Aerial Vehicle (UAV) technology, -also known as drones- has raised concerns on the safety of critical locations such as governmental buildings, nuclear stations, crowded places etc. Computer vision based approach for detecting these threats seems as a viable solution due to various advantages. We envision an autonomous drone detection and tracking system for the protection of strategic locations. It has been reported numerous times that, one of the main challenges for aerial object recognition with computer vision is discriminating birds from the targets. In this work, we have used 2-dimensional scale, rotation and translation invariant Generic Fourier Descriptor (GFD) features and classified targets as a drone or bird by a neural network. For the training of this system, a large dataset composed of birds and drones is gathered from open sources. We have achieved up to 85.3% overall correct classification rate.

Introduction

The volume of the Unmanned Aerial Vehicle (UAV) -also known as *drone*- industry expands constantly, making these gadgets accessible to more and more ordinary citizens, with cheaper prices. This situation enforces authorities to change the security paradigm for strategic locations such as nuclear power stations, touristic hot spots, governmental buildings etc.. Drones can easily be converted to dangerous weapons by loading them with explosives. A couple of terrorist attack attempts involving drones have been reported [1]. In addition to this, more and more incidents on drones threating civil aviation have being reported [2]. Also, espionage and privacy issues are other problems [3]. They cannot be detected efficiently with conventional methods, such as RADARs etc. due to their size and small electromagnetic signatures [4].

These developments have raised concerns as drones are immune the conventional defence systems, mainly developed for military purposes. They cannot be detected efficiently with conventional methods, such as RADARs etc. due to their size and small electromagnetic signatures [4]. Therefore, new kind of counter measures are being evaluated by industry and academia. Among these methods, RF signal detection aims to detect the RF signals between the operator and the drone. Another set of method rely on the X-band and micro-doppler RADAR technology. There exists also acoustics systems where the specific noise emitted from the drone rotors is detected. Finally, computer vision is a versatile choice in addition to these approaches [5][6][7].

All of these methods have their advantages and drawbacks, however computer vision approach distincts itself with its effectiveness due to its rich features and robustness [4]. In this work, we have used Generic Fourier Descriptor (GFD) to characterize the binary shapes, (e.g. silhouettes) of drones and birds, which is an effective algorithm proposed by [9]. These are two dimensional scale, translate and rotation invariant features, which are defining shapes with high detail [9]. One of the main challenges in computer vision approach is the difficulty of distinguishing birds from the drones [8]. Therefore, we have developed a system which compare the GFDs of binary silhouettes of birds and drones to classify.

Detecting Drones with a Computer Vision Approach

As mentioned previously, computer vision for detecting drones is a more robust, feasible and effective method compared to other existing ones. Object recognition with computer vision algorithms has been gaining popularity in recent years combined with the deep learning framework. Convolutional Neural Networks (CNNs) are already accepted as the state-of-the-art for object identification by many [10]. It is a deep learning technique, which autonomously learns the optimal features for classification by imagery, thus does not depend on human crafted features [11].

We also started to see certain examples of application of CNNs for drone detection such as [12][13][14], using ordinary daylight camera. CNNs may started be considered as the most recent and state-of-the-art solution in the literature, however they require extensive computational cost, especially for training phase. Their real-time operability with cost-effective material is not feasible due to computation burden. In addition, their accuracy may be still low for certain circumstances such as low resolution, insufficient dataset etc. For all these different variations of the object image such as different illumination, different poses etc., the dataset shall be fed by high number of training examples.

BoW approach is known as an effective computer vision algorithm for general object classification problem, with applications such as [16],[17],[18]. Rather than CNNs, [15] uses SURF based keypoint features of grayscale drone, bird and background image patches. The authors propose a new kind of extended bag-of-words (BoW) approach for classification. In this article, we propose a GFD based approach for classifying image patches composed of birds and drones similar to those in [15].

Generic Fourier Descriptors

Fourier Descriptors have been used as an efficient shape descriptor [19]. The distances of each contour pixel to the center of mass of the 2D object silhouettes is represented as a vector. Fourier Transform of this vector gives a unique description of the shape as the transform itself is shift, scale and rotation invariant.

Generally, the lower frequencies of the transform contains more information on the major structural parts of the object. If we interpret the mechanism of the algorithm, we can state that higher frequencies of the transform correspond to the more inten-

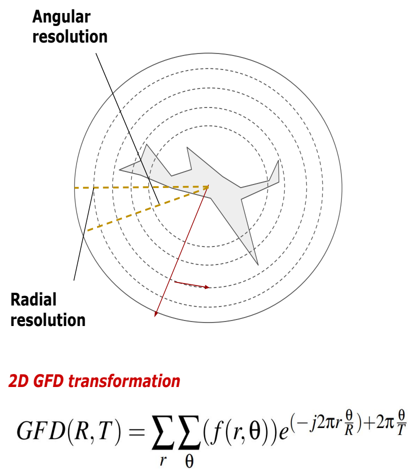

Figure 1. *The GFD transformation of a 2D object silhouette.*

sive ripples on the contour.

However, even this approach can differentiate non-similar silhouettes with high efficiency, the classification performance degrades as the contours get similar. In addition to this, as mentioned previously, this algorithm only considers the shape of the outside contours. However, the form of the silhouettes can contain very important and distinctive information such as holes etc. An approach taking into account the complete silhouette shall be more robust to noise which can miss certain number of pixels.

Generic Fourier Descriptor (GFD) is a method proposed by [9], which takes into account the 2D object silhouette in contrast the unidimensional Fourier Descriptors. The idea is to first raster and transform the pixels of the silhouette to polar coordinates with chosen angular and radial resolutions. Normalized 2D Fourier transform (Eq. 1) of this rastered function generates two dimensional matrix which we use as the representation of the shape. When this result is being used for classification with various algorithms, it is vectorized [9].

$$GFD(R,T) = \sum_{r}\sum_{\theta}(f(r,\theta))e^{(-j2\pi r\frac{\theta}{R})+2\pi\frac{\theta}{T}} \quad (1)$$

As it is normalized, this method is intrinsically scale invariant. And due to polar mapping by taking the center of mass as the origin, it is also translate and rotation invariant just like the regular Fourier Descriptors. Fig. 1 illustrates the GFD calculation of a 2D object silhouette.

Detecting Drones with GFD

Thanks to its proven effectiveness and strong scalability, invariability properties of GFDs of [9], we have aimed to use it for our purpose of detecting drones, while preventing false alarms caused by birds. We have imagined a system as illustrated in Fig. 2, where a steady wide-angle daylight camera observes the horizon to detect flying objects by applying a background subtraction algorithm. These background subtraction algorithm may be chosen from wide options such as Gaussian Mixture Models [20], median filtering etc. [21]. Note that, this procedure can be performed by thermal cameras for nocturnal operation, as binary silhouettes are used. Then, after normalizing and centering the silhouette, we calculate GFD according to (1). We have chosen 16 radial resolutions and 9 angular resolutions for calculation (therefore, we have feature vectors composed of 144 scalar values for each object.).

Another advantage of GFD compared to regular Fourier Descriptor is the fact that grayscale intensity value of each pixel can be used. This property can add an effective dimension for distinction, especially for thermal cameras. However, in this paperwork, we have not used this property, but just the binary silhouettes.

In order to train our system, we have created a dataset, composed of images of flying birds and drones, which are acquired from open sources. To seperate the object pixels from the background, a special image segmentation algorithm is applied. We have composed the dataset from the images, where the object is darker than the background (i.e. sky). The images are chosen to be relatively low resolution in order to reflect the target case, where the autonomous tracker detect a small flying object in wide angle. All images are converted to gray scale and rescaled to 64x64 pixels. Fig. 3 shows few of the images from the dataset both for drones and birds, where we have acquired silhouettes.

Region Growing algorithm is chosen as the image segmentation algorithm to seperate the object silhouettes from the background in the images due to its efficiency [22]. Region Growing is a method, where pixel neighborhoods are evaluated in an iterative manner, starting from an initial seed point. Over course of the algorithm, the pixels are defined as background or foreground, by applying a clustering criterion. Note that, this procedure is just for creating a large silhouette database. Note that, for the few cases where the object is brighter than the background, we have inversed illumination.

In order not to lose any shape information, we have not applied any morphological operations after image segmentation phase. For most of the cases, after image segmentation there is no need for further processing to acquire the true binary silhouettes of the objects. However, in case there are more than one disconnected pixel groups, the algorithm chooses the largest pixel group as the true binary silhouette of the object. To apply GFD, the binary silhouettes of the objects have to be centered in a 2D plane, where their center of mass is the origin. Note that, as GFD is a scale and rotation invariant transform, there is no need for rescaling or rotation for the silhouettes. Fourier Transformation results are normalized and reshifted before further processing. Following this, we have created a neural network composed of approximately 10000 neurons to classify GFD features in to birds and drones.

We have used 410 drone images and 930 bird images. A 5-fold approach is followed, where 4/5 of the samples are always used for training and the 1/5 of the samples is used for testing. In addition to this, to assure the validity of the experiments, the training samples are again divided in an additional 5-fold manner, where the test group is used for developing regularization parameters during the optimization of the neural network.

We have acquired an overall 85.8% accuracy on the test groups. The Table 1 shows the confusion matrix for a test group. As it can be seen, the GFD based algorithm is especially effective at detecting bird shapes, with a true rate of 93.7%. However, we

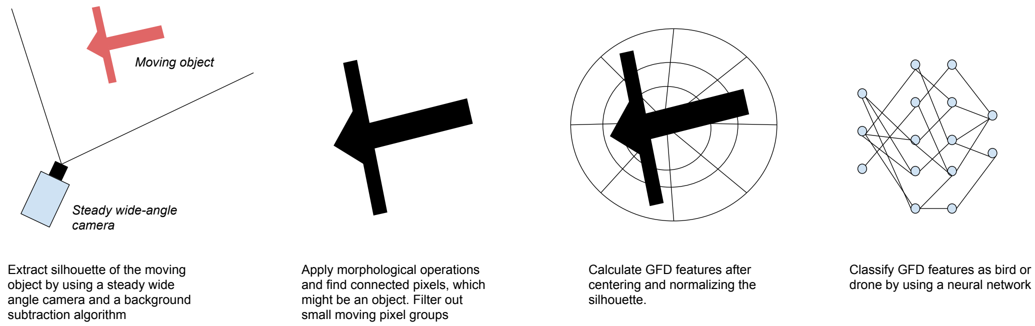

Figure 2. *A few examples of the 64x64 grayscale bird, drone and background patches from the images we have collected from the internet.*

Figure 3. *A few examples of the 64x64 grayscale bird and drone patches from the images we have collected from the open sources, in order to create silhouette database.*

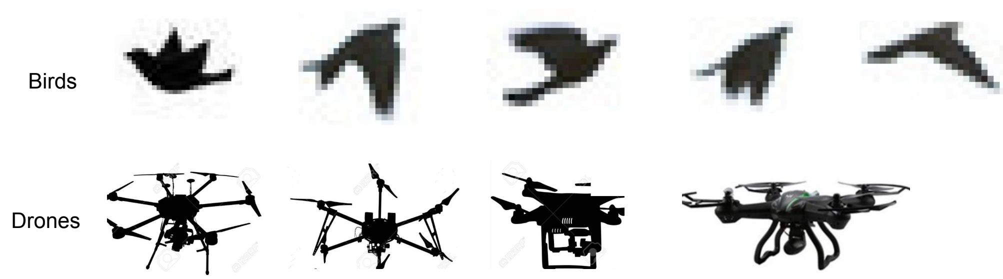

Figure 4. *A few examples of the 64x64 grayscale bird and drone patches from the images we have collected from the open sources, in order to create silhouette database.*

see that only 64.6% of drones are correctly identified. The overall accuracy shows that GFD is a novel 2D shape descriptor for discriminating bird and drone silhouettes for an autonomous drone surveillence system.

We have also tested our approach for the case where we have used the graphical grayscale silhouettes of the drones and birds, acquired from the internet. Therefore, we were not obliged to use region growing algorithm for segmentation. The same processes are applied for this dataset also. Fig. 4 shows some examples from the graphical silhouettes of birds and drones which are taken from [23] for the limited access version. Similarly for the previous case, we have calculated the confusion matrix for the test groups as in Table 2. A total of 540 samples are used, whom 419 are bird silhouettes and the rest are drone silhouettes. As for the previous case, 80% of the dataset is used for training and the rest for the test, in a 5-fold manner. We have acquired an overall accuracy of 77.1% on the test group.

Performance Comparison to a CNN with Smaller Datasets

In this context, where there are numerous type and models of drones exist as potential intruders, it may be difficult to gather sizable amount of training images for each of them. Not only model and type, but high number of color options make the case even harder. For instance, gathering hundreds of images for a specific type, model and color of a drone under various poses, orientations under various illumination, background etc. may be virtually impossible. Hence we have tested the performance of the proposed approach with a much smaller dataset and compared the performance to a convolutional neural network. In addition to this, we have combined the resulted scores with the feature vector of the GFD approach, in order to inspect whether there is a performance gain with a possible complementary effect.

For this purpose, we have combined a total of 80 training drone images of similar type and 40 training bird images. There are 29 test drone images and 13 test bird images. The system uses the same infrastructure have build a CNN architecture with 3 convolutional layers, where the first one is composed of 15, second one composed of 20 and last one composed of 25 5x5 filters. At the end, there is a fully connected layer with 256 neurons. We have observed that, the CNN attains a 85.08% of accuracy and the proposed GFD approach attains a 93.10% accuracy on test dataset, as it can be seen in Table 3 and Table 4.

Next, we have first calculated the score of the CNN architecture, then added into the GFD feature vector before neural network classification. As it can be seen in Table 5, for the same small test dataset, the accuracy increases to 100%. This further implies that CNN and GFD can operate in a complementary nature for increasing the performance.

Conclusion

As it can be seen in this paperwork, GFD can be chosen as shape descriptor tool for an autonomous drone surveillence system. We believe the performance presented in this work can be augmented with larger datasets. This approach may compete with the currently preferred CNN based algorithms, which require extensive computational power and very large datasets, while providing no information on the curvature of the object. Another advantage of GFD based algorithm is its potential ability to be used for motion based changes. This can be performed by analyzing temporal changes of GFD features. For instance, we can guess that movements of a drone shall be much more stable compared to a bird, especially considering the wing flapping. For CNNs this analysis is cumbersome and needs other type of deep learning techniques, which takes into account the spatio-temporal relations.

At first, we have combined the GFD based shape features with a neural network to test its accuracy in our context of detecting flying drones and discriminating birds with a steady camera for a dataset composed of real images or artificial models. We have attained a 85.8% and 77.1% accuracy for these datasets, respectively. Next, we have tested the performance of the proposed approach with a much smaller dataset and compared it to a 3-layer CNN architecture. As expected, CNN which requires large datasets performed worse compared to the proposed GFD approach. However, we have shown that the best performance can be gained by combining CNN scores with GFD feature vectors.

Table 1. Confusion matrix for the classification of bird and drone silhouettes in test set.

Output Class	**Bird**	**207** **68.3%**	**29** **9.6%**	*87.7%* *12.3%*
	Drone	**14** **4.6%**	**53** **17.5%**	*79.1%* *20.9%*
		93.7% *6.3%*	*64.6%* *35.4%*	**85.8%** **14.2%**
Target Class		**Bird**	**Drone**	

Table 2. Confusion matrix for the classification of bird and drone silhouettes directly taken from the artificial graphical drawings.

Output Class	**Bird**	**20** **18.3%**	**12** **11.0%**	*62.5%* *37.5%*
	Drone	**13** **11.9%**	**64** **28.7%**	*83.1%* *16.9%*
		60.6% *39.4%*	*84.2%* *15.8%*	**77.1%** **22.9%**
Target Class		**Bird**	**Drone**	

Table 3. Confusion matrix for the CNN classification of bird and drone images with smaller dataset.

Output Class	**Bird**	**22** **52.38%**	**1** **2.38%**	*95.65%* *4.35%*
	Drone	**7** **16.66%**	**12** **28.57%**	*63.15%* *36.85%*
		75.86% *24.14%*	*94.3%* *5.7%*	**85.08%** **14.92%**
Target Class		**Bird**	**Drone**	

Table 4. Confusion matrix for the GFD classification of bird and drone silhouettes generated from images with smaller dataset.

Output Class	**Bird**	**25** **59.52%**	**0** **0%**	*100.0%* *0.00%*
	Drone	**4** **9.56%**	**13** **30.95%**	*68.42%* *31.58%*
		86.20% *13.80%*	*100.0%* *0.00%*	**93.1%** **6.90%**
Target Class		**Bird**	**Drone**	

Table 5. Confusion matrix for the combined CNN+GFD feature classification with smaller dataset.

Output Class	**Bird**	**29** **69.04%**	**0** **0%**	*100.0%* *0.00%*
	Drone	**0** **9.56%**	**13** **30.95%**	*100.0%* *0.00%*
		100.0% *0.00%*	*100.0%* *0.00%*	**100.0%** **0.00%**
Target Class		**Bird**	**Drone**	

References

[1] Humphreys, T. O. D. D. ”Statement on the security threat posed by unmanned aerial systems and possible countermeasures.” Oversight and Management Efficiency Subcommittee, Homeland Security Committee, Washington, DC, US House (2015).

[2] Wild, Graham, John Murray, and Glenn Baxter. ”Exploring civil drone accidents and incidents to help prevent potential air disasters.” Aerospace 3.3 (2016): 22.

[3] Villasenor, John. ”Observations from above: unmanned aircraft systems and privacy.” Harv. JL & Pub. Pol'y 36 (2013): 457.

[4] Peacock, Matthew, and Michael N. Johnstone. ”Towards detection and control of civilian unmanned aerial vehicles.” (2013).

[5] Yoon, Ji Hwan, Hao Xu, and Luis R. Garcia Carrillo. ”Advanced Doppler radar physiological sensing technique for drone detection.” SPIE Defense+ Security. International Society for Optics and Photonics, 2017.

[6] Hearing, Brian, and John Franklin. ”Drone detection and classification methods and apparatus.” U.S. Patent No. 9,275,645. 1 Mar. 2016.

[7] Solodov, Alexander, et al. ”Analyzing the threat of unmanned aerial vehicles (UAV) to nuclear facilities.” Security Journal (2017): 1-20.

[8] Gke, Fatih, et al. ”Vision-based detection and distance estimation of micro unmanned aerial vehicles.” Sensors 15.9 (2015): 23805-23846.

[9] Zhang, Dengsheng, and Guojun Lu. ”Shape-based image retrieval using generic Fourier descriptor.” Signal Processing: Image Communication 17.10 (2002): 825-848.

[10] Ciresan, Dan C., et al. ”Flexible, high performance convolutional neural networks for image classification.” IJCAI Proceedings-International Joint Conference on Artificial Intelligence. Vol. 22. No. 1. 2011.

[11] Simard, Patrice Y., David Steinkraus, and John C. Platt. ”Best practices for convolutional neural networks applied to visual document analysis.” ICDAR. Vol. 3. 2003.

[12] Schumann, Arne, et al. ”Deep cross-domain flying object classification for robust UAV detection.” Advanced Video and Signal Based Surveillance (AVSS), 2017 14th IEEE International Conference on. IEEE, 2017.

[13] Saqib, Muhammad, et al. ”A study on detecting drones using deep convolutional neural networks.” Advanced Video and Signal Based Surveillance (AVSS), 2017 14th IEEE International Conference on. IEEE, 2017.

[14] Aker, Cemal, and Sinan Kalkan. ”Using deep networks for drone detection.” Advanced Video and Signal Based Surveillance (AVSS), 2017 14th IEEE International Conference on. IEEE, 2017.

[15] Unlu, Eren, Emmanuel Zenou, and Nicolas Riviere. ”Ordered minimum distance bag-of-words approach for aerial object identification.” Advanced Video and Signal Based Surveillance (AVSS), 2017 14th IEEE International Conference on. IEEE, 2017.

[16] Yang, Yi, and Shawn Newsam. ”Bag-of-visual-words and spatial extensions for land-use classification.” Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems. ACM, 2010.

[17] O'Hara, Stephen, and Bruce A. Draper. ”Introduction to the bag of features paradigm for image classification and retrieval.” arXiv preprint arXiv:1101.3354 (2011).

[18] Bu, Fengju. ”Human Behavior Recognition based on Conditional Random Field and Bag-Of-Visual Words Semantic Model.” International Journal of Signal Processing and Pattern Recognition 8.1 (2015): 23-32.

[19] Persoon, Eric, and King-Sun Fu. ”Shape discrimination using Fourier descriptors.” IEEE Transactions on systems, man, and cybernetics 7.3 (1977): 170-179.

[20] Lee, Dar-Shyang. ”Effective Gaussian mixture learning for video background subtraction.” IEEE transactions on pattern analysis and machine intelligence 27.5 (2005): 827-832.

[21] Piccardi, Massimo. ”Background subtraction techniques: a review.” Systems, man and cybernetics, 2004 IEEE international conference on. Vol. 4. IEEE, 2004.

[22] Adams, Rolf, and Leanne Bischof. ”Seeded region growing.” IEEE Transactions on pattern analysis and machine intelligence 16.6 (1994): 641-647.

[23] [image] Available at: https://fr.123rf.com/ [Accessed 10 Sep. 2017].

Author Biography

Dr. Eren Unlu has graduated from Bilkent University, Turkey in 2011 in electronics engineering and achieved his Master's degree on mobile telecommunications from Institut EURECOM, France in 2013. Following this, he has acquired his PhD in 2016 from CentraleSupelec, France where he examined optimal bandwidth allocation algorithms for an OFDMA based on-chip wired RF interconnect. Dr. Unlu has worked in CentraleSupelec in 2016 as a post-doctoral researcher in data science and machine learning for a year. Currently, he is still continuing his research as a post-doctorate in ISAE-SUPAERO, Toulouse, France on computer vision and machine learning.