Automatic Classification of UAVs with a Conventional Radar-Based Surveillance System

Gianpaolo Pinelli
Ingegneria dei Sistemi S.p.A, 24, Via Enrica Calabresi, 56124, Pisa, ITALY

g.pinelli@idscorporation.com

ABSTRACT

This paper proposes an innovative approach to classify mini-UAVs (Unmanned Aerial Vehicle) using a conventional counter-drone radar-based aerial surveillance system. State-of-the-art machine learning (ML) techniques applies for distinguishing UAV tracks from other kind of aerial (birds, airplanes, helicopters, etc.) and ground (car, pedestrian, etc.) targets. Experimental results obtained on a meaningful real data set shows the capability to classify UAVs with accuracy higher that 98%. Moreover, a second stage of classification, suitable to separate fixed wing (FW) from rotary wing (RW) UAVs allows us to reach accuracy performance higher that 93%.

Keywords: *Surveillance Radar, Drone, Counter Unmanned Aerial Vehicle, Classification, Support Vector Machines*

1.0 INTRODUCTION

Nowadays UAVs make it possible to imagine many previously unavailable and non-cost-effective applications, such as safeguarding human life, environmental monitoring, and last but not least, security and defense. UAVs have become essential for military applications but their use is currently increasing for hostile activities, making the drone one of the main threats to face on the battlefield as shown in the Ukraine theater.

In recent years, research and development (R&D) activities of counter-UAV technologies and systems have multiplied. In fact, it is worth noting that conventional aerial surveillance systems are not enabled to face effectively such specific threats, often summarized by the acronym LLS, i.e. capable of flying at low altitude (Low), light (Light), and small (Small). IDS BK (*Black Knight*) belongs to this class of experience, being a multi-sensor (EO/IR/Radar) aerial surveillance system specifically designed to deal with UAVs.

IDS BK is a mid-range C-UAV system capable of detecting, tracking, and, if needed, neutralizing (jamming, spoofing) potential UAV threats to critical infrastructures and sensitive public and private areas, up to 5Km for mini-UAVs@-3dBsm Radar Cross Section (RCS). This paper proposes the recent R&D improvements of the radar sensor capabilities, capable to classify UAVs in an automatic fashion up to the maximum distance of detection, without interrupting the normal search and detection of threats present in the three-dimensional spherical observation space, i.e. by using a conventional surveillance radar system.

A state-of-the-art feature-based ML algorithm is used to distinguish in an automatic fashion UAV from other kinds of aerial and ground targets such as manned platforms (e.g. airplanes, helicopters) and natural false alarms (e.g. birds), along the entire coherent integration time (CIT). On wide collections of real data, it has been derived an accuracy higher than 98%. Moreover, the ability to distinguish the wing type can be shown (e.g. Rotary Wing (RW) vs. Fixed Wing (FW)), reaching an accuracy higher than 93%.

In the following §0, we describe the radar signal processing chain necessary to detect and track multiple targets at the same time, the available pieces of information which can be exploited for the classification, the radar configurations and the measurement campaign performed to train the classifier. In §0, we describe the machine learning techniques adopted to design the classifier and evaluate its performance. In §0 we show the experimental results obtained on the available data.

2.0 RADAR SIGNAL PROCESSING

This work aims to design a technique able to distinguish drones from other objects, and FW from RW UAVs, using the signal received by a surveillance X-band radar working with a *Linearly Frequency Modulated Continuous Wave* (LFMCW) transmitted waveform.

A *surveillance* radar operates with a rotating antenna to discover, detect and track multiple targets at the same time [1]-[2]. As a surveillance radar constantly seeks the space to find new targets, the target is continuously observed by the radar, usually with very small (*Time on Target* (ToT)), in the order of 10 msec.

The most widely used radar-based architecture for classification and identification of drones is the *tracking* one, which illuminates a single target for a fairly longer time, in the order of 1 s. The tracking radar holds the antenna in the direction of the designed target, and allows the analysis of features describing the intrinsic movements of the target through the analysis of the time variations of the Fourier spectra of received signals, which is called the micro-Doppler analysis [5]-[6].

In counter-drone application, the necessity to detect and track multiple targets at the same time can only be met by a surveillance radar. For this reason, only techniques to classify the received radar signal from a surveillance radar can be applied.

2.1 Pre-Processing

The radar processing chain to detect and track multiple targets is shown in Figure 1. Using the same notation of [1]-[2]-[6], the received raw radar signal can be seen as a matrix of values defined in the *Fast Time (FT)/Slow Time (ST)* domain. FT samples identify the range sampling of the received echo (sweep) for a specific azimuth location, while ST samples identify the azimuth coordinates corresponding to the consequential transmitted radar pulses. Each sample is a complex value, identified by the I (In-phase) and Q (Quadrature) received channels.

High pass filter, 1D (along the FT direction) Fast Fourier Transform (FFT) and a calibration procedure is applied to the raw signal to obtain a Range Profile Matrix (RPM) of Radar Cross Section (RCS) values in the Range / Slow Time domain [1]. This is the first piece of information, which can be used for drone classification. However, we resort to the anomalies detected by the classical radar signal processing chain of Figure 1. The RPM are processed with 1D FFT along the Doppler direction, by taking a number of slow time samples, which identifies the *Coherent Integration Time* (CIT) [1]. In our case, the CIT is of around 30 ms, and it identifies a radar azimuth cell of 4.5°. A high pass filter is also used to kill the zero-velocity components, removing the stationary clutter from the radar detections.

Finally, the Range Doppler Matrix (RDM) is obtained, which represents the RCS w.r.t. Range and Doppler (or, equivalently, radial velocity [6]) dimensions. The RDM could be directly fed to a Neural Network (NN) based algorithm for drone classification, as in [7]. In our application, RDM is used to find the local maxima points with fairly high RCS values. Those points identify the targets and are called *Detections*. The row and column of the Detection identify respectively the Range and the radial Velocity of the target.

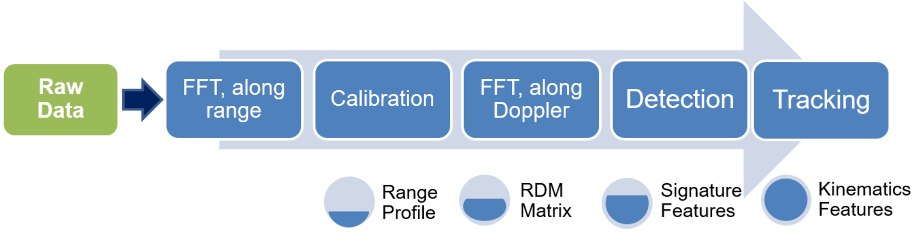


Figure 1: Radar signal processing to detect and track multiple targets, and available pieces of information which can be exploited for drone classification.

For each Detection, the RDM is used to define a set of signature features. For example, the Detection amplitude describes the RCS, and the ratio between the amplitude of the Detection vs the mean amplitude of pixels in the same row describes the *Signal to Noise Ratio* (SNR). Several other ratios can be defined in the RDM, describing maximum and average amplitudes between regions around the Detection and its surroundings [3]-[4]-[7].

Then, Detections are clustered using the Range/Azimuth/Radial Velocity domain. Two or more Detections very close to each other in all the three domains are grouped together in a *cluster* called *Plot* [7], and assigned to the same observation [10].

Finally, using a Kalman filter [9] designed to work in Range/Azimuth domain, the Plots are associated to one or more *Tracks* describing the trajectory and the velocity of the targets, and to predict their future positions. Tracks can be used to evaluate kinematic features of the target, and are useful for classification.

2.2 Features Definition

Here in after, we define *Detections*, *Plots* and *Tracks* (Figure 2), according to the same notation shown in [7]. For each Detection, a set of descriptors derived from RDM has been defined, including RCS and SNR. Detections are clustered into Plots, and the number of Detections in the Plot, also constitutes an useful feature which can be used for classification.

The Plots are observations for the tracking algorithm, which groups them into tracks describing the trajectory of the target. This allows the definition of the kinematic features. They can be evaluated by considering a *Segment of track*, which is a part of the track obtained after a fixed number of observations. Each observation of Kalman filter (i.e. each Plot) can be obtained after one full antenna rotation, which is the time after which the radar antenna will be again in the direction of the target.

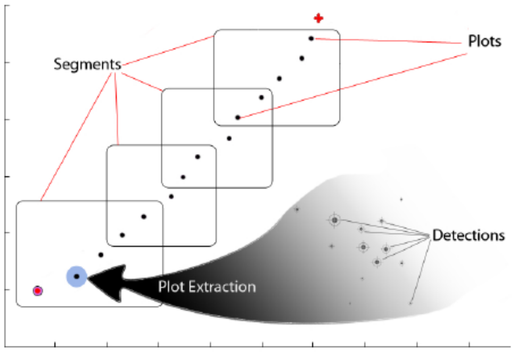


Figure 2: Detections clustered into Plots, and Plots are used to define the track of the target. A Segment of track is a set of a fixed number of Plots in a track [7].

In this study, the classification performance has been analyzed w.r.t. the length of the Segment of track, in the range (4 – 10). We call this parameter NTREF.

In each Segment of track, a set of kinematic features [7] is defined to describe the target trajectory in the segment. For the NTREF Plots in the Segment, the mean and the standard deviation of the above mentioned signature features are considered. The total number of features for each Segment of Track is 50, of which 30 are signature-based and 20 kinematic-based.

2.3 Radar Parameters and Configurations

The radar operates in the X-band (9.35 GHz), with a Transmitted Power of 4 W and a LFMCW. Its Bandwidth can vary with the configuration, with a maximum of 100 MHz. It performs a 2D scan in Range/Azimuth domain. The central elevation angle has to be set by the operator, and the antenna elevation beam is 22.8°. The Pulse Repetition Frequency (PRF) is set to 3.3 KHz.

Table 1: Radar parameters used by the three main configurations.

Configuration	**Operative_600m**	**Operative_2km**	**Operative_4km**
Parameter	**Value**		
Signal Start Frequency	9.3625 GHz	9.3625 GHz	9.3625 GHz
Band	75 MHz	18.867 MHz	9.75 MHz
Antenna rounds per minute	20	20	20
Max Range	624 m	2100 m	4200 m
Range Resolution	2 m	7.95 m	15.9 m
Max Target Speed	96 Km / h	96 Km / h	96 Km / h
Pulse Repetition Frequency	3339 Hz	3339 Hz	3339 Hz
Samples in a Sweep	624	624	624

In the measurement campaigns, mainly three radar configurations have been considered, each characterized by the maximum range of the radar: 624 m, 2.1 Km and 4.2 Km. They are named respectively *Operative_600m*, *Operative_2km* and *Operative_4km*.

Table 1 summarizes the main parameters of the three radar configurations. The classification algorithms have been designed for each configuration separately. Finally, a classification algorithm has been trained for all the data in all the configurations.

2.4 Measurement Campaign

The measurement campaign was performed by acquiring a set of UAVs:

- Commercial RW as Phantom3 Pro (DJI), Typhoon 4K (Yuneec), Jetson (NVIDIA), Bebop2 (Parrot),
- Commercial FW as Disco (Parrot),
- IDS RW as FlySmart 2.0, Colibrì, Nik, FlyNovex,
- IDS FW as FlyFast, FlySecur.

When possible, GPS position and time of the drone flight was saved, to help the necessary labelling process of the radar signal. A semi-automatic procedure was developed to label radar data whether GPS information of the target is available or not.

A very high number of non-drone objects were recorded during the measurement campaign. They were non-cooperative targets, such as birds, airplanes, cars, helicopters, walking people. Even without GPS information, in many cases, it was possible to label them as "false alarms" (FA) using the knowledge of the position of the drone during the acquisition.

Table 2 shows the number of acquired samples for FA and Drone classes, and Table 3 for FW and RW classes, w.r.t. configuration and NTREF parameter. Of course, the higher NTREF is, the lower the number of samples to train the classifier. The number of recorded samples of the FA class is much higher for the 2km and 4km configurations than for the 600m one, because the space exploited by the radar is much bigger.

Table 2: Number of acquired samples for Drone vs FA classification, for each configuration and values of the number of antenna rotations to define a Segment of track.

	NTREF	**4**	**6**	**8**	**10**
Operative_600m	**FA**	8914	5512	3419	2200
	Drone	2463	2007	1658	1364
Operative_2km	**FA**	81824	56822	41446	31525
	Drone	3243	2819	2454	2170
Operative_4km	**FA**	40833	30128	23423	18837
	Drone	1470	1258	1078	918

Table 3: Number of samples for Fixed Wing vs Rotary Wing discrimination.

	NTREF	**4**	**6**	**8**	**10**
Operative_600m	FW	689	539	419	328
	RW	1774	1468	1239	1036
Operative_2km	FW	1496	1275	1097	961
	RW	1747	1544	1357	1209
Operative_4km	FW	398	303	225	153
	RW	1072	955	853	765

3.0 CLASSIFICATION ALGORITHM

3.1 Training Process

Given an object under test, the purpose of the algorithm is to decide whether the object is a Drone or not, and if it is a Drone, to distinguish between FW and RW Drone.

Many classification algorithms have been compared to this purpose, not only in terms of performance, but also in terms of computational time for training, and overfitting avoidance. Classical algorithms from Machine Learning (ML) theory [12] have been taken into consideration, including KNN (K Nearest Neighbors), Adaboost, Gradient boost, Support Vector Machines (SVM) and Multi-Layer Perceptron.

The comparison between different ML techniques goes beyond the purpose of this paper. We choose to use SVM with radial basis kernel [11], because it obtained an acceptable trade-off between performance, training time and overfitting avoidance.

The training process of the SVM classifier has been performed trough the *s*-fold cross validation scheme [12]. The samples acquired during the measurement campaign have been split into *s* subsets. Each subset includes a number of samples such that the ratio between samples from different classes is the same as in the original dataset. The samples from the same acquisition are always included into the same subset. Moreover, the training always performs on samples coming from different acquisitions w.r.t. the ones in which it is tested. This allows to design a more robust classifier, and to give a more trustful estimate of the classification performance, and thus to predict its behavior when dealing with new samples.

During the *s*-fold cross validation process, the hyper-parameters and the subset of features within the 50 are chosen. The optimal subset of features would be given by the exhaustive search of all possible combinations of k features, with $1 \leq k \leq 50$, which is not feasible with standard hardware resources. For this reason, we adopted a suboptimal search with the *Sequential Floating Forward Feature Selection* method [12], which proved to be a very good trade-off between computational time for training and performance.

The radial basis kernel SVM needs the definition of the two hyper-parameters C and γ [11]. They have been searched choosing the best obtained by three different methods: the exhaustive search on a custom grid, the Newton-Bayes search [12] and the Automatic Model Selection method from Chapelle described in [12].

The suboptimal subset of features and the suboptimal set of SVM hyper-parameters have been searched by optimizing the classifier *accuracy*.

We choose to design two SVMs: the first to distinguish between Drone and FA, the second between FW and RW. Thus, we defined a two-stage SVM classifier. In our analyses, this approach has proven to be better in terms of performance and robustness w.r.t. the direct classification between the three classes FA / RW Drone / FW Drone.

3.2 Performance Evaluation

The performance of the classifier has been evaluated in terms of *accuracy*, per-class *recall* and *precision* indexes, which are defined as in [12].

During the *s*-fold cross validation process, we evaluated the mean performance (i.e. accuracy, recall, precision) among each of the *s* classification experiments.

A small number of acquisitions were hidden to the classifier during the training process (i.e. *holdout* [12] – 10% of the available data). This allows also to measure the performance of the classifier with a final blind test which allows to check its robustness.

After the *s*-fold and the blind test, all available data are re-split into new subsets, and the performance is re-evaluated. This is another precaution taken against overfitting.

Each classifier performance has been thus evaluated in three versions: the one obtained during s-fold training process, the one obtained during blind test, and the one obtained in the final tests re-partitioning the available data. The mean of the three has been called *Global Index* (GI), and can refer to all indexes (accuracy, recall, precision). In this paper, results are presented only in terms of GI.

Generally speaking, the higher the GI, the better the classifier. The comparison of the three indexes obtained during s-fold, blind test, and final re-partitioning, allows to check the robustness of the classifier. Generally speaking, the more similar the three indexes, the more robust the classifier.

The performance are presented w.r.t. the radar configuration and the number NTREF of antenna rotations to define a segment of track. We also show the performance of a classifier designed for all the radar configurations.

4.0 EXPERIMENTAL RESULTS

In the following, we present the experimental results obtained from the SVM classifiers with the data acquired during the measurement campaign. The following tables list the mean accuracy for each configuration, and for NTREF = 4, 6, 8. Table 4 lists the performance for the Drone vs FA classification, and Table 5 for FW vs RW. Figure 3 shows the trend of the accuracy for each configuration, w.r.t. NTREF in the range (4 – 10). All indexes are expressed in terms of GI.

Table 4 shows that all Drone vs FA classifiers have good performance, while our comparison among the three accuracies show that those classifiers are also robust. Accuracy is around 98% for the *2km* and *4km* Configurations, and for the classifier trained for all the configurations (*All_Conf*). Accuracy is lower, around 95%, for the *600m* configuration, but this does not mean that the overall radar performance is worst in that configuration. In fact, the *2km* and *4km* configurations are characterized by a much higher number of false alarms than the *600m* one. Those FAs are generally well classified by the algorithm, leading to a higher accuracy.

As a matter of fact, we observe that if we analyze the performance only for the samples belonging to the Drone class, the classification algorithm for the *600m* Configuration achieves the best performance. The most likely error committed by the classifiers is the "missed detection", i.e. samples from Drone class erroneously assigned to FA class, and it is more likely to occur as the range increases.

The *All_Conf* classifier shows that accuracy is higher than 98%, meaning that less than two segments of tracks out of 100 are misclassified. Figure 3 shows that the performance of Drone / FA classifier is not afflicted by the choice of the number of antenna rotations to define a segment of track. In this case, it is preferable to use the lowest value, i.e. NTREF = 4, it is not necessary to gather more information waiting for further antenna rotations. The classifier proves also robustness for each NTREF parameter.

Table 4: Accuracy (GI) obtained for Drone vs FA classification, w.r.t. Radar Configuration and NTREF parameter. Accuracy is expressed in percentage. For each classifier the table lists the accuracies obtained during s-fold training process, blind test and final GI.

Drone / FA		**Accuracy %**		
	NTREF	4	6	8
Operative_600m	**s-fold**	95.79	95.93	95.98
	Blind	89.63	91.83	89.52
	GI	95.46	95.40	95.62
Operative_2km	**s-fold**	98.87	98.73	98.73

Drone / FA		Accuracy %		
	Blind	99.40	99.32	99.29
	GI	98.82	98.79	98.74
Operative_4km	**s-fold**	97.88	97.59	97.48
	Blind	99.17	98.17	98.87
	GI	98.32	97.69	97.99
All_Conf	**s-fold**	98.28	98.27	98.33
	Blind	99.40	99.37	99.44
	GI	98.29	98.35	98.35

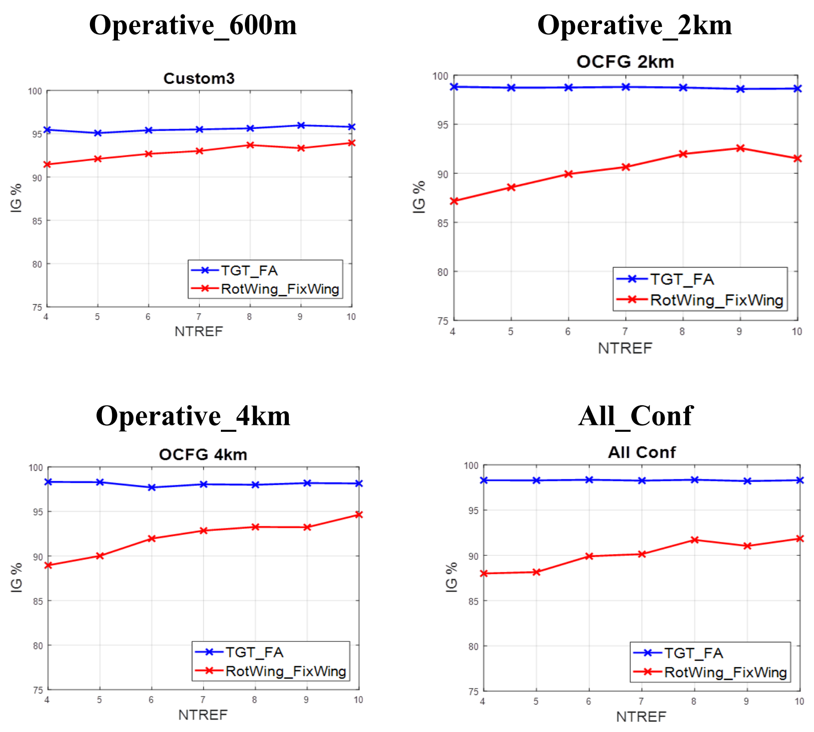
The FW vs RW classifier, instead, can take advantage of using more antenna rotations to improve both performance and robustness, as shown in Table 5 and in Figure 3. Accuracy is around 88-90% for NTREF = 4, and improves to 92-94% for NTREF = 10.

Performance of FW / RW classifiers are good generally speaking, but not as good as the ones obtained by the Drone / FA classifiers. Due to the lower number of samples from FW class, the RW class is generally better classified.

Finally, we believe that the performance of both FW / RW and Drone / FA classifiers could improve by increasing the number of samples for the Drone class in the database. Table 5: Accuracy (GI) obtained for Fixed Wing vs Rotary wing vs Rotary wing classification.

FW / RW		Accuracy %		
	NTREF	4	6	8
Operative_600m	**s-fold**	91.35	92.56	93.59
	Blind	93.94	95.15	96.12
	GI	91.43	92.67	93.69
Operative_2km	**s-fold**	86.83	91.04	92.93
	Blind	94.33	81.75	85.95
	GI	87.17	89.92	91.97
Operative_4km	**s-fold**	90.39	94.09	95.58
	Blind	80.64	89.88	85.42
	GI	88.96	91.96	93.26
All_Conf	**s-fold**	87.88	90.24	91.97
	Blind	93.97	86.95	85.68
	GI	88.00	89.90	91.71

Figure 3: Mean accuracy (GI) vs NTREF parameter, for each radar configuration.



5.0 CONCLUSIONS

In this paper, we have shown a novel radar processing algorithms designed to classify UAV versus non-UAV tracks and, within the UAV class, to discriminate among RW versus FW drone type. The multi-stage classification here proposed adheres to the SVM architecture and it is based on a proper selection of identifying signature and kinematics features.

The different stages of classification have been trained through extensive UAV measurement campaigns conducted in a controlled environment, using X-band LFMCW IDS surveillance radar with different radar parameter settings, target and scenarios. Experimental results are highly promising, showing drone/no drone average correct classification accuracy around 98% for the 2Km and 4Km radar configuration and FW/RW accuracy around 92-94% taking advantage from collection of data acquired by higher antenna rotations (NTREF=10).

6.0 REFERENCES

[1] Tait P., “Introduction to Radar Target Recognition”, IET, 10.1049/PBRA018E, 2005

[2] Sullivan R. J., “Radar Foundations for Imaging and Advanced Concepts”, Electromagnetics and Radar, Scitech Publishing; Revised ed. edition (June 30, 2004)

[3] Ghadaki H., Dizaji R., “Target Track Classification for Airport Surveillance Radar (ASR)”, *2006 IEEE Conference on Radar*, 2006.

[4] Dizaji R., Ghadaki H., "Classification System for Radar and Sonar Applications", Patent US7 567 203.

[5] De Wit J. J. M., Harmanny R. I. A. and Molchanov P., "Radar micro-Doppler feature extraction using the Singular Value Decomposition," *2014 International Radar Conference*, Lille, 2014, pp. 1-6.

[6] Chen V., "The Micro-Doppler Effect in Radar" – *Artech House Radar Library*, 2012.

[7] Mohajerin N., Histon J., Dizaji R. and Waslander S. L., "Feature extraction and radar track classification for detecting UAVs in civillian airspace," *2014 IEEE Radar Conference*, Cincinnati, OH, 2014, pp. 0674-0679.

[8] Vojtech M., "Objects identification in signal processing of FMCW radar for Advanced Driver Assintance Systems", Diploma Thesis Assignment, Czech Technical University in Prague, Faculty of Electrical Engineering

[9] Blackman S., Popoli R., "Design and Analysis of Modern Tracking Systems", Boston MA, Artech House, 1999.

[10] Klaasing K., Wollher D., Buss M., "A Clustering Method for Efficient Segmentation of 3D Laser Data", *2008 IEEE International Conference on Robotics and Automation*, Pasadena CA, May 2008, pp. 4043-4048.

[11] Ivanciuc O., "Applications of Support Vector Machines in Chemistry", *Reviews in Computational Chemistry, Volume 23*, 2007.

[12] Guyon, Gunn S., Nikravesh M., Zadeh L. A., "Feature Extraction. Foundations and Applications", Springer, 2006.