Paper

Adaptive procedure for automatic modulation recognition

Ferdinand Liedtke

Abstract—An adaptive procedure for automatic modulation recognition is described. With it the automatic modulation classification and recognition of radio communication signals with a priori unknown parameters is possible effectively. The results of modulation recognition are important in the context of radio monitoring or electronic support measurements. The special features of the procedure are the possibility to adapt it dynamically to nearly all modulation types, and the capability to recognize continuous phase modulation (CPM) signals like Gaussian minimum-shift keying (GMSK) too. A time synchronization to the symbol rate is not necessary.

Keywords— modulation recognition, modulation classification, signal analysis, radio monitoring.

1. Introduction

The results of modulation recognition are important in the context of radio monitoring or electronic support measurements. Originally the research was conducted for short wave or high frequency (HF) communication signals in the radio frequency (RF) range between 1 and 30 MHz. In this range there exist many different signal modulation types or wave forms, and often the systems or signal characteristics are not standardized. Furthermore, modern HF radio systems are able to change their parameters and their modulation types continuously, dependent on the quality of the transmission channel. The developed modulation recognition procedure can be used not only for the analysis of HF signals, but also for radio communication signals from higher RF ranges, e.g., VHF or UHF. In those ranges lots of military and civil mobile radios are running. Though many of those systems are standardized with known signal characteristics, it will be of interest to detect the activity of specific systems and their currently used modulation types. The additional challenge in the higher frequency ranges are the often used CPM waveforms, e.g., GMSK with varying B*T values. For those soft-keyed wave forms it is more difficult to find relevant signal values suited for modulation recognition than for wave forms with hard keying. A modulation recogniser developed for digitally modulated signals with hard keying is described in [1]. In this context it is important to remember that in general the symbol rate or the time points for symbol synchronization are not known for the non-authorized receiver. In many papers about modulation recognition the a priori knowledge of the exact values for centre frequency, bandwidth, and symbol rate of every interesting signal is assumed, e.g., [2, 3]. This will be appropriate for an application like the adjustment of a software radio which expects several well defined signal wave forms. In contrary to that, this presupposition cannot be maintained in general for applications like radio monitoring or electronic support measurements. For these applications a special robustness against parameter inaccuracies is necessary. A further challenge for the modulation recognition described here is the quantity of different wave forms, several of which are often totally unknown from the beginning. Therefore it is desirable to have a recognition procedure which adapts easily to the various wave forms.

For the future, some further development of the recognition procedure is planned.

2. Principle of the procedure

The necessary pre-processing of the received signal is represented in Fig. 1. From the down converted and digitised signal a spectrum is computed with the aim to accomplish a *spectral segmentation*, i.e., to estimate the centre frequency f_c and the bandwidth B of a significant spectrum part.

With the determined parameters the signal is appropriately shifted in frequency and filtered to get the complex base band signal z. On this occasion the segmentation procedure is not discussed in detail. The signal z is fed into the *modulation recognition* module. The essential parts of this module are represented in Fig. 2. The figure parts marked in grey designate the differences to the formerly used modulation recogniser [1]. The further processing is performed in three branches: The upper two branches with the *squaring of absolute values*, the *squaring of complex values*, and the following digital Fourier transforms (DFTs) are used for exploitation of the symbol rate information, which perhaps may be included in the signal, by means of the *spectral line detection*. The *squaring of absolute values* could reveal the appropriate spectral line, indicating the symbol rate for digitally modulated signals with hard keying while the *squaring of complex values* is provided for CPM. Within the described modulation recogniser the symbol rate estimation is not necessary for obtaining a time synchronization, the sole aim is the estimation of the approximate symbol dwell time T. Time T is needed for the calculation of the appropriate values for the difference phase $D\varphi$ within the *feature extraction* module in the third branch. In future, the spectral symbol rate information, which differs in dependence on the modulation type, could perhaps be utilized for the classification process too, but in the currently used recogniser it is not yet used for this purpose. Before the *feature extraction* is carried out, a *coordinate resolving* module provides the signal in polar coordinates a and φ.

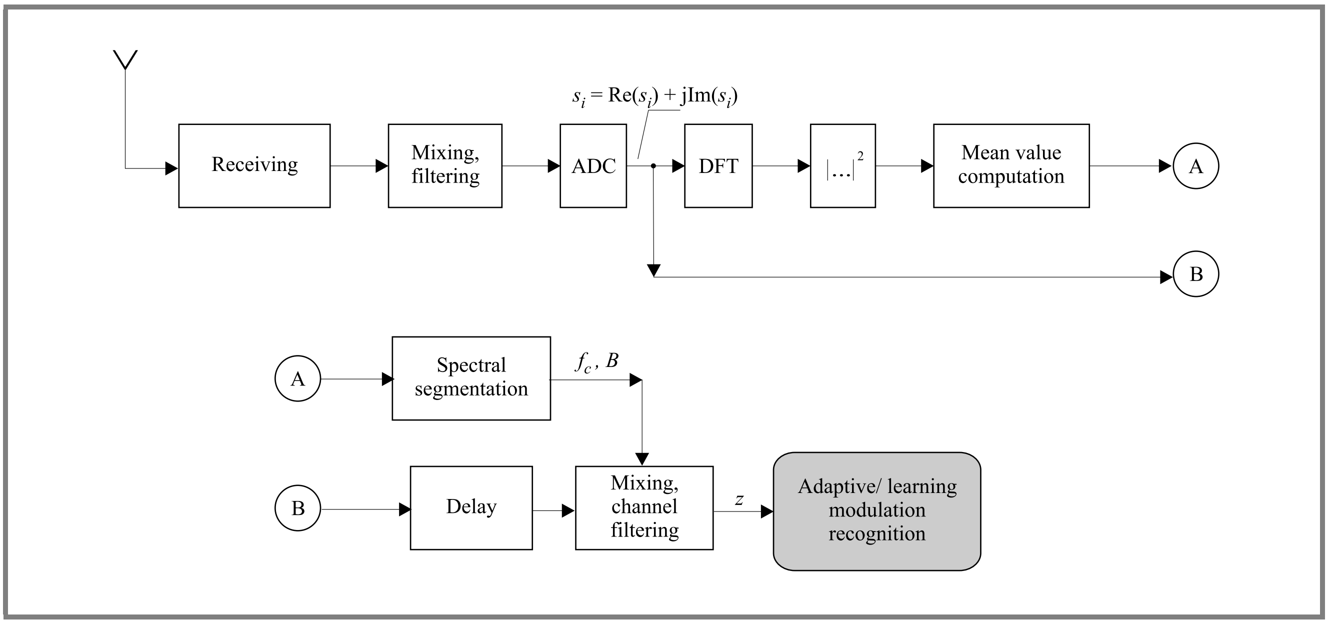

Fig. 1. Signal analysis with modulation recognition.

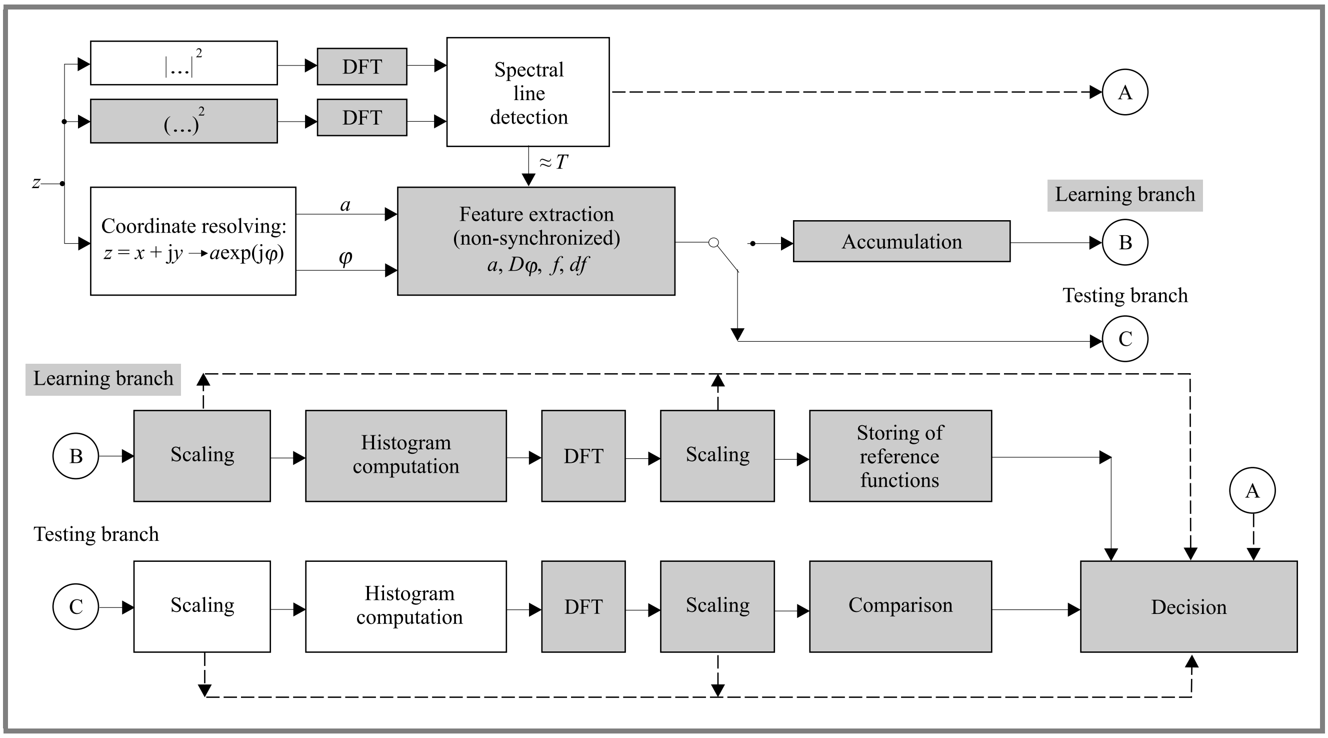

Fig. 2. Adaptive modulation recognition, non-synchronized.

The *feature extraction* module provides relevant values of the following signal parameters: amplitude a, difference phase $D\varphi$, instantaneous frequency f, and derivation of the instantaneous frequency df. Concerning the parameter φ, difference values and not the phase values themselves are used for the purpose to be resistant to a non-exact frequency tuning.

The time difference for extracting every pair of phases for calculating $D\varphi$ is T. The parameter f is essential for recognizing simple frequency modulated signals like frequency-shift keying and the derivative of f is provided for chirp signals. The procedure for extracting the relevant parameter values is discussed in the last paragraph of this section.

After *feature extraction* there is a splitting into two branches, a *learning branch* and a *testing branch*. By means of the *learning branch*, it is possible to adapt to the various modulation types or wave forms. For this aim the relevant parameter values are accumulated in a learning phase and further processed as follows: the parameter values are appropriately scaled and written into *histograms*, one histogram for each of the four parameters. The histograms are then transformed into a picture domain by means of the DFT. The transformed functions or picture functions are a kind of characteristic functions with the property that the interesting information is concentrated in the first part of each function. It was found out that the use of the first quarter of the respective picture function values is sufficient. In the following text this part of a picture function is still called "picture function". The picture functions are appropriately *scaled* and *stored as reference functions*, for every modulation type a set of four functions. After having finished the learning of all interesting wave forms, the change-over switch following the *feature extraction* module (see Fig. 2) is switched into the lower position and the testing or working phase begins. The signal processing in the testing branch is the same as that in the learning branch up to the *scaling* of the picture functions. In the following *comparison* the actual set of picture functions are compared to all stored function sets and the set with the least deviation determines the modulation type. The deviation measure used here is the least mean square (LMS) and the fusion of the results obtained from the processing of the individual parameters is done by simply adding the LMS results. The dashed lines in Fig. 2 refer to the intention to search (in future) for additional features, obtained from spectral or scaling parameters.

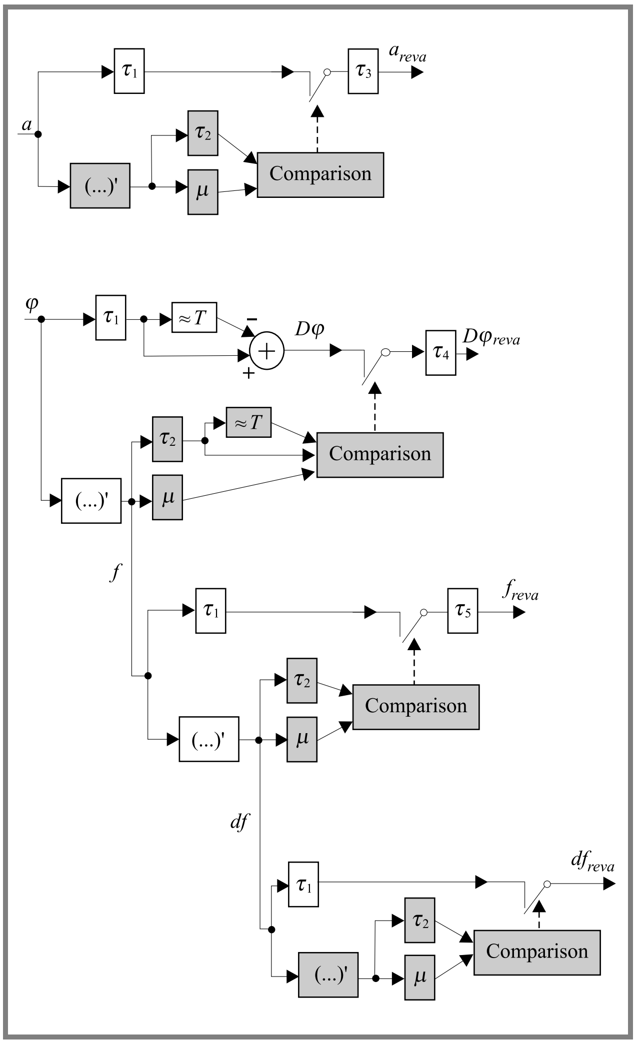

Fig. 3. Feature extraction, non-synchronized.

The extraction of the relevant parameter values for a, $D\varphi$, f, and df is represented in Fig. 3. Because the extraction was not intended to need the synchronization to the symbol rate, another method had to be worked out. The **re**levant **va**lues are those where the eyes of their so-called eye patterns are wide open, i.e., the parameters have reached their steady states. The way to find these values is to continuously compute the *time derivative* of each parameter (in Fig. 3 depicted with (...)'), to compute the *sliding mean* μ of these derived values, and to continuously perform a *comparison* of μ with the corresponding original parameter value. It was found out that the relevant values are those which are near to the mean of the derived values. This is a similar procedure as the search for extreme values of a function by means of its derivative. Here, the *comparison* is not carried out with the derivative value zero, but with the mean of the derived values. The intention is to compensate parameter trends, which normally are still existent. Concerning the phase it has to be remembered that always difference phase values have to be considered.

3. Experimental results

The whole procedure including signal generation, anti-aliasing filtering, noise addition, reception filtering, and modulation recognition was developed and tested on a PC with the software tools of MATLAB. Successful tests were also carried out with real signals obtained from the mobile radio systems GSM, TETRA, and TETRAPOL.

A class space of 15 modulation types was defined, most of them digitally modulated wave forms. The frequently used modulation types amplitude-shift keying with two states (ASK2), frequency-shift keying (FSK2), minimum-shift keying (MSK) including different GMSK forms, and phase-shift keying (PSK) with different numbers and positions of relevant phase states are considered among others. The class space was supplemented with a white Gaussian noise class (WGN) and a reject class (REJ). All classes are represented in Table 1. The listed signal to noise ratios (SNRs) are explained in the next paragraph. The arrangement with altogether 17 classes was chosen to challenge the recognition procedure. In realistic scenarios the class number could be reduced in many cases, which alleviates the classification task.

To learn the class specific reference functions 16 learning experiments per class were carried out, whereby for every experiment a signal segment with a length corresponding to 256 symbols was used. The number of individual signal samples per symbol dwell time T was chosen to be 8. In the next step the appropriate decision levels were learned

Table 1
Used classes/wave forms

No.	Name of class/wave form	Comments	SNR_{start} [dB]
0	REJ	Reject class	—
1	WGN	White Gaussian noise	—
2	ASK2	Amplitude-shift keying, 2 states	15
3	FSK2	Frequency-shift keying, 2 states, $\beta = 0.85$	12
4	MSK	Minimum-shift keying	12
5	GMSK05	Gaussian MSK, B*T = 0.5	12
6	GMSK03	Gaussian MSK, B*T = 0.3	12
7	PSK2	Phase-shift keying, 2 states	9
8	$\pi/2$ DPSK2	Differential PSK2, keying with $+/-\pi/2$	9
9	PSK4	Phase-shift keying, 4 states	14
10	$\pi/4$ DPSK4	Differential PSK4, keying with $+/-\pi/4$	14
11	OPSK4	Offset PSK4	14
12	PSK8	Phase-shift keying, 8 states	20
13	ASK2/PSK8	Hybrid digital modulation type, 2 amplitude states for each of the 8 phase states	25
14	CLOVER	ASK2/PSK8 with Chebyshev pulse weighting	25
15	LICHIRP	Linear chirp	20
16	CHIRPKEY	Chirp keying with linear up and down chirp	20

with 16 additional experiments. The used SNRs are related to those values with which an authorized receiver could reach a symbol error probability of 10^{-4} with non-coherent or differential demodulation, whereby a coding gain was not considered. These SNRs differ in dependence on the modulation type and they are called SNR_{start}. The use of different SNR_{start} values for the different modulation classes was realized because the authorized receiver expects different SNRs for different modulation types too, compare Table 1. For example the SNR_{start} value for the PSK4 types is 14 dB and that for the PSK2 types is 9 dB. A quarter of the experiments was carried out with SNR_{start}, the next quarter was performed with $SNR_{start} - 2$ dB, the third quarter has $SNR_{start} - 4$ dB and the last quarter was performed with $SNR_{start} - 6$ dB, i.e., the mean SNR was $SNR_{start} - 3$ dB. In that way it was possible to take into account different SNRs with comparatively low experiment numbers. The testing or working phase was carried out with 64 experiments per class, each experiment with a signal length corresponding to 256 symbols too. The principle of choosing the SNRs was the same as that used in the learning phase.

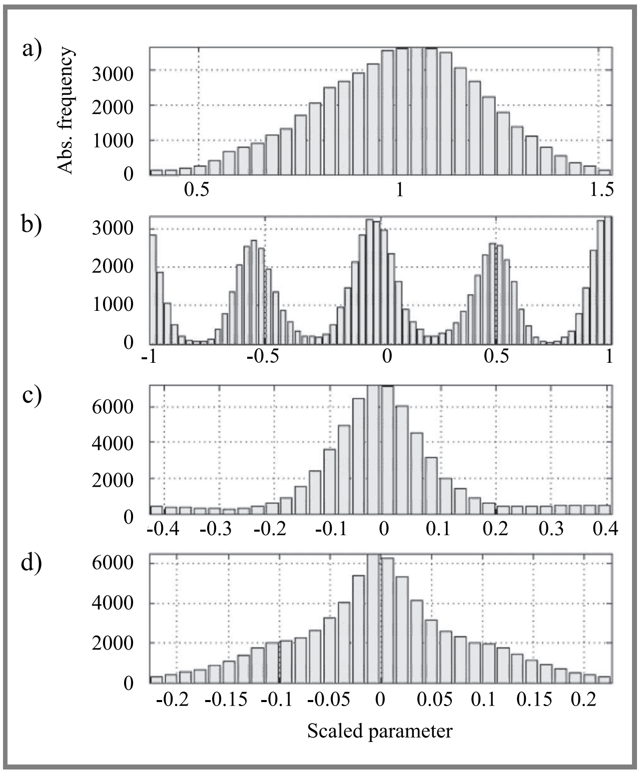

Fig. 4. Parameter histograms for PSK4: (a) amplitude a; (b) difference phase $D\varphi$; (c) instantaneous frequence f; (d) differential instantaneous frequency df.

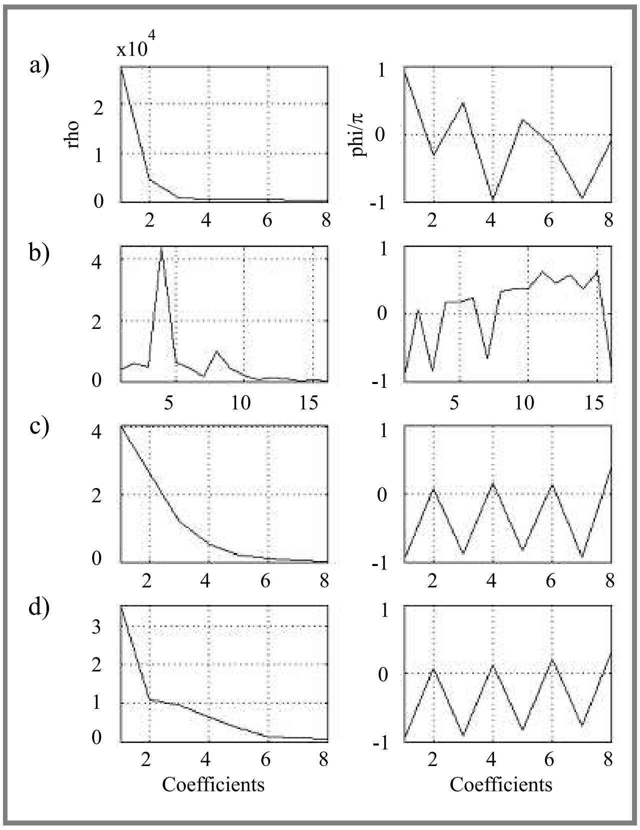

Fig. 5. Picture functions for PSK4: (a) amplitude a; (b) difference phase $D\varphi$; (c) instantaneous frequence f; (d) differential instantaneous frequency df.

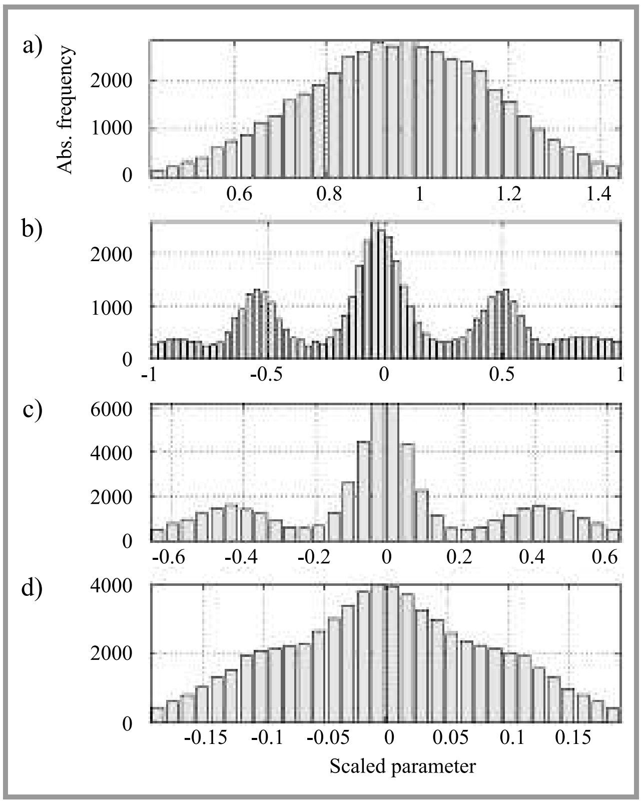

Fig. 6. Parameter histograms for OPSK4: (a) amplitude a; (b) difference phase $D\varphi$; (c) instantaneous frequence f; (d) differential instantaneous frequency df.

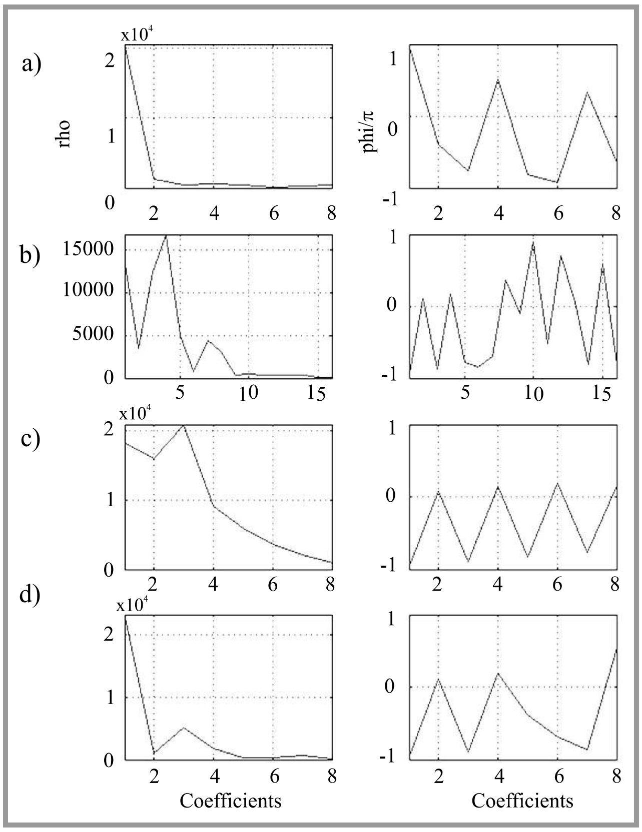

Fig. 7. Picture functions for OPSK4: (a) amplitude a; (b) difference phase $D\varphi$; (c) instantaneous frequence f; (d) differential instantaneous frequency df.

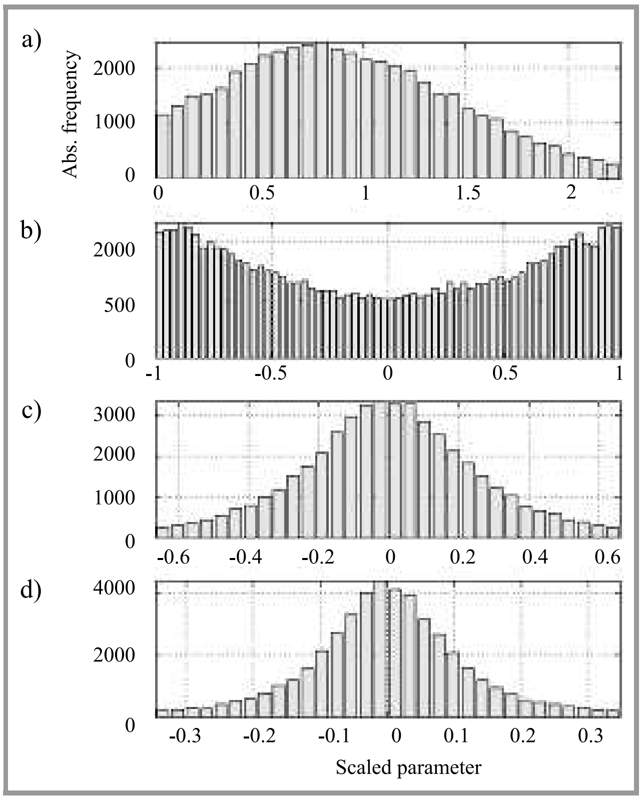

Fig. 8. Parameter histograms for WGN: (a) amplitude a; (b) difference phase $D\varphi$; (c) instantaneous frequence f; (d) differential instantaneous frequency df.

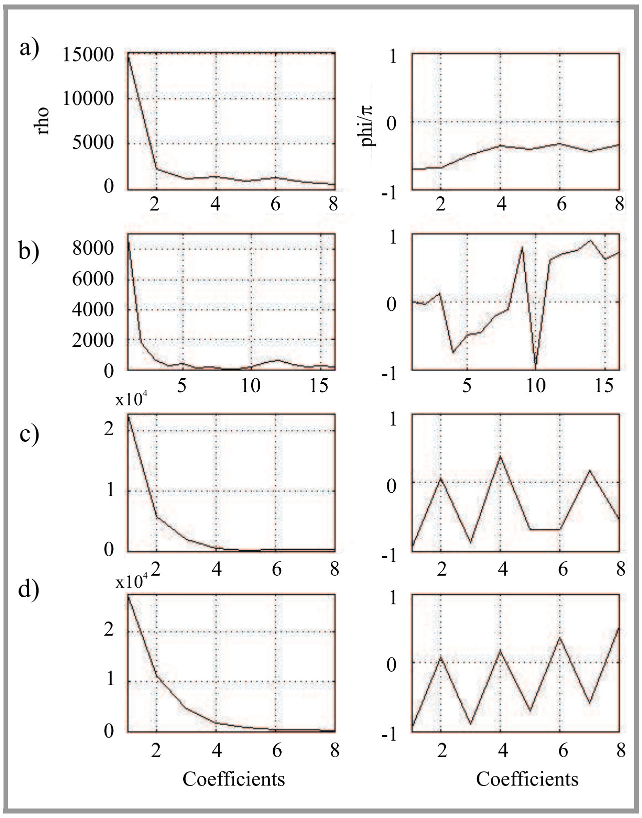

Fig. 9. Picture functions for WGN: (a) amplitude a; (b) difference phase $D\varphi$; (c) instantaneous frequence f; (d) differential instantaneous frequency df.

Table 2
Confusion matrix, $\mathrm{SNR}_{\mathrm{start}} - 6\ \mathrm{dB} \leq \mathrm{SNR} \leq \mathrm{SNR}_{\mathrm{start}}$

Input	Output class																
	ASK2	FSK2	MSK	GMSK05	GMSK03	PSK2	$\pi/2$ DPSK2	PSK4	$\pi/4$ DPSK4	OPSK4	PSK8	ASK2/PSK8	CLOVER	LICHIRP	CHIRKEY	WGN	REJ
ASK2	98.4																1.6
FSK2		100															
MSK			**70.3**	**23.4**	**4.7**												1.6
GMSK05			**32.8**	**54.7**	**12.5**												
GMSK03				**21.9**	**75**												3.1
PSK2						98.4											1.6
$\pi/2$ DPSK2							100										
PSK4								89.1									10.9
$\pi/4$ DPSK4									98.4								1.6
OPSK4							3.1			87.5							9.4
PSK8											96.9						3.1
ASK2/PSK8												96.9					3.1
CLOVER													95.3				4.7
LICHIRP														95.3			4.7
CHIRKEY															100		
WGN																85.9	14.1

Table 3
Confusion matrix, $\mathrm{SNR}_{\mathrm{start}} - 9\ \mathrm{dB} \leq \mathrm{SNR} \leq \mathrm{SNR}_{\mathrm{start}} - 3\ \mathrm{dB}$

Input	Output class																
	ASK2	FSK2	MSK	GMSK05	GMSK03	PSK2	$\pi/2$ DPSK2	PSK4	$\pi/4$ DPSK4	OPSK4	PSK8	ASK2/PSK8	CLOVER	LICHIRP	CHIRKEY	WGN	REJ
ASK2	67.2																32.8
FSK2		84.4														14.1	1.6
MSK			**39.1**	**26.6**	**32.8**												1.6
GMSK05			**6.3**	**40.6**	**51.6**												1.6
GMSK03				**17.9**	**82.8**												
PSK2						68.8										4.7	26.6
$\pi/2$ DPSK2							78.1									6.3	15.6
PSK4						1.6	1.6	64.1								1.6	31.3
$\pi/4$ DPSK4									78.1								21.9
OPSK4						9.4			3.1	60.9							26.6
PSK8										4.7	60.9						34.4
ASK2/PSK8												71.9					28.1
CLOVER													64.1				35.9
LICHIRP				4.7	3.1				1.6					78.1			12.5
CHIRKEY					34.4										64.1		1.6
WGN																85.9	14.1

As a first demonstration example, histograms of the four examined parameters a, $D\varphi$, f, and df for a PSK4 signal are depicted in Fig. 4. The results are obtained from 16 experiments with 256 symbols per experiment, i.e., from 4096 symbols altogether. This comparatively high symbol number was chosen to clearly depict the typical characteristics. For a normal recognition process the exploitation of 256 symbols is entirely sufficient. It can be seen from Fig. 4 that the $D\varphi$ histogram has 4 peaks. This is the main characteristic indicating the PSK4 wave form.

The picture functions are depicted in Fig. 5. In the left column the results for the absolute values rho and in the right column the phase results phi$/\pi$ of the histogram transforms are depicted. The important result is the peak position at the abscissa value 4 in Fig. 5 on the left hand side, second row. The corresponding results of an OPSK4 signal are depicted in Figs. 6 and 7. The peaks in the $D\varphi$ histogram (see Fig. 6) are less evident than those of the corresponding PSK4 histogram. This becomes noticeable in Fig. 7, left hand side, second row, too. However, for sufficiently good SNRs, OPSK4 can be classified correctly. For lower SNRs some classifications as PSK2 or $\pi/2$ DPSK2 are possible. For comparison with the WGN situation the corresponding results for WGN are depicted in Figs. 8 and 9. It can be observed that neither the histograms nor the picture functions give any indication of a digitally modulated signal (with two or more characteristic parameter positions). Therefore it could be understood that the WGN class can be easily discriminated from the other classes. The principal considerations concerning the other classes, not yet discussed, are similar.

The final classification results are represented in form of so-called confusion matrices (see Tables 2 and 3). The names of all treated classes were entered into the left column and into the uppermost row. The left column depicts the classes fed into the modulation recogniser and the classification results were entered into the columns with the corresponding output class names indicated in the uppermost row. For an ideal recogniser all results, each 100%, are contained in the diagonal matrix elements. These elements are indicative of correct classifications.

The learning of the reference picture functions was performed with 16 experiments per class. For every experiment a signal segment length corresponding to 256 symbols was used. Another 16 experiments per class were carried out to learn the necessary variances of the deviations of the test picture functions from the stored reference functions. These values were needed for arranging the appropriate decision levels. The SNRs were chosen in the range $\mathrm{SNR}_{\mathrm{start}} - 6$ dB $\leq$ SNR $\leq \mathrm{SNR}_{\mathrm{start}}$ as discussed above. The testing phase consisted of 64 experiments for each class. The results depicted in the confusion matrices are indicated in percent related to 64. The difference between Tables 2 and 3 is the different SNR selection in the test experiments. While the tests for Table 2 were performed with $\mathrm{SNR}_{\mathrm{start}} - 6$ dB $\leq$ SNR $\leq \mathrm{SNR}_{\mathrm{start}}$ the tests for Table 3 were carried out with SNRs 3 dB worse. The results in Table 2 show that most of the classification results are contained in the diagonal elements or in their neighbourhood. The results for the CPM signals MSK, GMSK05, and GMSK03 are marked with bold numbers to indicate the close relationship between the three modulation classes. In other words, the spread of classification results is caused by the similarity of the wave forms and could not really assessed as false classifications. Most of the non-successfully classified signals were assigned to the reject class REJ. This is a desired result because a rejection or a classification as WGN is preferred to a false classification. The classification of two OPSK4 signals (3.1%) as $\pi/2$ DPSK2 is caused by experiments with comparatively low SNR. Both wave forms are similar in certain aspects.

The classification results presented in Table 3 are worse because the SNRs were 3 dB lower. The false classification of 34.4% of the CHIRKEY signals as GMSK03 is not really critical because CHIRKEY is a here synthesized artificial signal wave form, which was chosen very similar to GMSK03. The false classification of 9.4% of the OPSK4 signals as PSK2 is not surprising because for the smaller SNRs these wave forms are also similar. The numbers of the other false classifications are still comparatively moderate.

For assessing the discussed experimental results one has to realize that for real scenarios the class number could be reduced in many cases. This alleviates the classification task. For the future, further work is planned on the following items: Test and integration of additional classification features, appropriate evaluation and fusion of the feature parameter values, adaptation of the classification procedure, and tests.

4. Conclusions

An automatic modulation recognition procedure is described which is suited for many modern signal wave forms including CPM. The automatic modulation recognition is interesting in the context of radio monitoring or electronic support measurements. For this application the signal parameters are often not known a priori, e.g., the centre frequency, the bandwidth, and the symbol rate. As a typical pattern recognition procedure the modulation recognition has a learning and a testing phase. During the learning phase an adaptation to the new wave forms is easily possible. The procedure was tested with 16 wave forms with different SNRs. The results are comparatively good. For the future some further improvements are planned.

References

[1] F. Liedtke, "Computer simulation of an automatic classification procedure for digitally modulated signals with unknown parameters", *Sig. Proc.*, vol. 6, no. 4, pp. 311–323, 1984.

[2] J. Palicot and C. Roland, "A new concept for wireless reconfigurable receivers", *IEEE COM Mag.*, pp. 124–132, July 2003.

[3] M. L. D. Wong and A. K. Nandi, "Automatic digital modulation recognition using artificial neural network and genetic algorithm", *Sig. Proc.*, vol. 84, pp. 351–365, 2004.

Ferdinand Liedtke is working in the department for Radio Communications and Electronics (FE) of the Research Institute for Communications, Information Processing, and Ergonomics (FKIE) of the FGAN in Germany. He is working on the topics detection, segmentation, classification, and automatic modulation recognition of signals with a priori unknown parameters for many years. He has written papers and got patents. Several of the modulation recognisers offered by national and international companies are based on the concepts worked out in the FGAN-FKIE-FE. Dr.-Ing. Liedtke is further interested in modern radio systems and their susceptibilities.

e-mail: liedtke@fgan.de

FGAN-FKIE-FE

Neuenahrer Str. 20

D-53343 Wachtberg-Werthhoven, Germany