# Evaluation of features for the automatic recognition of OFDM signals in monitoring or cognitive receivers

Ferdinand Liedtke and Ulrike Albers

**Abstract**—The automatic recognition of signal types is an important task of monitoring receivers and also cognitive receivers. Several modulation recognition or classification procedures exist for single channel signal types while a simple robust procedure for automatic recognition of OFDM signals is lacking because of its numerous frequency channels lying close together. The task considered in this paper is the discrimination between OFDM (or multi-channel) signals and other signal types. The number of frequency channels of the OFDM signals is assumed to be unknown a priori. So, together with the automatic OFDM detection the estimation of the number of frequency channels is treated. Several discrimination features have been examined and the most promising ones are described: measures of the variation, of the skewness, of the kurtosis, and of the specific picket-fence shape of the spectrum which is typical for many OFDM signals. For a number of real-world OFDM samples, recorded from the high frequency range, results are presented. An automatic discrimination from single channel or noise like signals is achieved and the number of system channels can be estimated.

**Keywords**— *OFDM signal recognition, discrimination features, cepstrum evaluation, estimation of frequency channel number.*

## 1. Introduction

Automatic recognition of signal types is an important task for monitoring receivers and also cognitive receivers. A monitoring receiver is a non-cooperative receiver used for radio reconnaissance. A cognitive receiver is a cooperative receiver belonging to a cognitive radio which will be a future advancement of a software defined radio. In both kinds of receivers the knowledge of the signal type is needed for further signal processing, such as synchronization, equalization, and demodulation. Several modulation recognition or classification procedures exist for single channel signal types, compare, e.g., [1, 2], while a simple robust procedure for automatic recognition of orthogonal frequency-division multiplexing (OFDM) signals with its numerous frequency channels lying close together is lacking. OFDM signals play an important role in modern communication systems like, e.g., the wireless LAN systems IEEE 802.11 a/g and IEEE 802.16 (WiMAX)

or broadcasting systems like DAB and DVB-T. They are also considered, together with MC-CDMA signals, as possible signal types for the fourth generation of mobile communication systems.

Furthermore, many new OFDM modems are used for professional application. These modems can be used together with conventional radio sets. As a consequence, the occurrence of this signal type on the air is expected not only in the provided frequency ranges, e.g., the 2.4 GHz or 5 GHz bands, but in the whole interesting radio frequency range, i.e., from high frequency (HF) over very high frequency (VHF) to ultra high frequency (UHF). The main advantages of OFDM signals are their effective utilization of a preset frequency bandwidth and their robustness to impairments of the transmission channel, especially frequency selective fading.

Disadvantages of OFDM signals are their great demands on amplifier linearity and the necessity to provide a high precision for time and frequency synchronization. To alleviate the synchronization, OFDM signals are transmitted in block form and, typically, every block is preceded by a guard interval in which delayed versions of multi-path signal parts of the respective preceding block are expected. In a cooperative receiver these guard intervals are processed in another way than the signal parts containing the information so that the undesired effects of multi-path reception can be minimized.

One of the most demanding steps in designing an automatic detection and classification procedure is to find appropriate features with which the target signal type can be discriminated from other signal types. In the case of OFDM signals as discussed here, the aim is to find and evaluate several features suited for discrimination of the complete signal with all used frequency channels and to avoid the necessity to handle individual channels in advance. Otherwise, attempting to achieve such a channel separation, a very precise synchronization of frequency and time would be necessary which is not available at this level of signal processing, especially for a non-cooperative monitoring receiver. So, the particular number of frequency channels of an observed OFDM signal should not be relevant for the discrimination features which have to be found. After an automatic detection of an OFDM signal however, the estimation of the number of frequency channels is desired.

In this paper the extraction and the evaluation of altogether seven discrimination features are described. Before extracting the features a certain preprocessing of the signal samples is necessary.

## 2. Signals and preprocessing

For the considerations below it is assumed that the detection of signal energy and the segmentation in time and frequency were done in advance and that the signal sample was down converted appropriately to the centre frequency zero, resampled and filtered according to that bandwidth value which resulted from the spectral segmentation process. The final sampling rate was chosen with an oversampling factor of four with respect to the significant signal bandwidth.
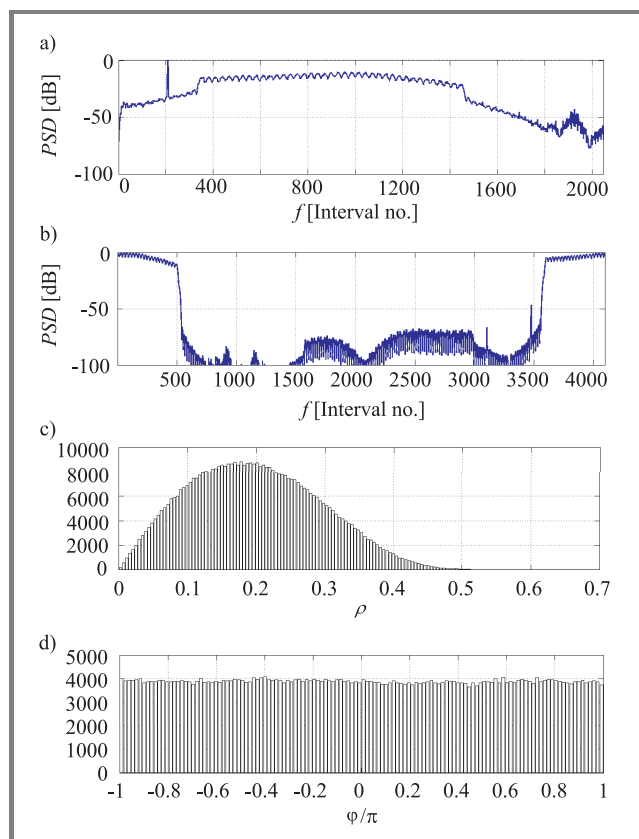


**Fig. 1.** A typical HF OFDM signal: (a) spectrum; (b) spectrum after preprocessing; (c) histogram of magnitude $\rho$, time domain; (d) histogram of phase $\varphi/\pi$, time domain.

To get an impression of the preprocessing, some characteristics of a typical HF OFDM signal with 39 frequency channels are shown in Fig. 1. In Fig. 1a the spectrum of the recorded real valued signal is depicted. As typical for OFDM signals with appreciable guard intervals, a picket-fence shape of the spectrum is observed. This shape is used to develop an efficient discrimination feature which will be described in Section 4. Figure 1b shows the spectrum after the preprocessing was completed. The signal is now complex valued. In Fig. 1c the histogram of the sig-

nal magnitude $\rho$ and in Fig. 1d the histogram of the phase $\varphi/\pi$ are depicted. The shapes of the histograms resemble those for white Gausian noise (WGN), i.e., a Rayleigh distribution for the magnitude and a uniform distribution for the phase. This is not surprising because the distribution of a superposition of many sine waves with equal amplitudes, equidistant frequencies, and different phases is approximately a Gaussian distribution. This fact will be utilized for the choice and the evaluation of the first six discrimination features which will be discussed in the next section.

Table 1
Considered signals

| Signal no. | Name | Remarks |
|---|---|---|
| 1 | QPSK$_\infty$ | Quaternary phase shift keying; $SNR = \infty$ |
| 2 | QPSK$_{14}$ | $SNR = 14$ dB |
| 3 | QPSK$_8$ | $SNR = 8$ dB |
| 4 | QPSK$_2$ | $SNR = 2$ dB |
| 5 | WGN | White Gaussian noise |
| 6 | ROC-39CH-MIL-1 | Rockwell modem; 39 channels; military version; sample 1 |
| 7 | ROC-39CH-MIL-2 | Rockwell modem; 39 channels; military version; sample 2; strong fading |
| 8 | NATO-39CH-1 | NATO modem; 39 channels; sample 1 |
| 9 | NATO-39CH-2 | NATO modem; 39 channels; sample 2 |
| 10 | BGR-39CH-TFC | Bulgaria; 39 channels; traffic |
| 11 | BGR-39CH-IDLE-1 | Bulgaria; 39 channels; idle (no traffic); sample 1 |
| 12 | BGR-39CH-IDLE-2 | Bulgaria; 39 channels; idle (no traffic); sample 2 |
| 13 | CHN-39CH | China; 39 channels |
| 14 | CZE-39CH | Czech Republic; 39 channels |
| 15 | MILST-39CH-S | Mil-standard 188-110a; 39 channels; short sample |
| 16 | RUS-45CH-TFC | Russia; 45 channels; traffic |
| 17 | RUS-60CH-TFC | Russia; 60 channels; traffic |
| 18 | MT-63CH | Multi-tone; 63 channels |
| 19 | NLD-64CH | Netherlands; 64 channels |
| 20 | RUS-93CH-TFC | Russia; 93 channels; traffic |
| 21 | RUS-93CH-IDLE | Russia; 93 channels; idle (no traffic) |

Table 1 shows a list of the considered signals. The signals no. 1 to 5 are synthetically generated ones which were

selected as typical **non-OFDM** signal types to verify the discrimination capability of the selected features described below. Signals 1 to 4 are single frequency channel quaternary phase shift keying (QPSK) signals with decreasing signal-to-noise ratios (SNRs) and signal 5 is a white Gaussian noise signal (WGN). The other signals are real-world OFDM samples recorded from the HF range.

These signals have different total bandwidths, different numbers of frequency channels, different symbol rates, different quality, and different sample lengths. The sample lengths vary from about 40,000 to about 100,000 (after resampling). These numbers seem large but their sizes have to be related to the over-sampling factor and to the number of frequency channels. The used oversampling factor is four. The frequency channel numbers range from 39 to 93. For the estimation of the spectrum details the number of symbols per frequency channel is important and the corresponding information content in the sample is only 1/(channel number) of the whole sample information. From a statistical point of view the different sample lengths are not satisfying, but, this fact corresponds to real scenarios and the signal processing has to cope with it. The signals are ordered according to increasing numbers of used frequency channels. Several signal types are represented with various samples which have different characteristics. For some signals also samples with idle mode (no information is transmitted) are included.

All signals were preprocessed as described above. The relevant bandwidth value of the QPSK signals was chosen as the symbol rate and the out of band spectral parts were filtered out with the same low pass filter which was used for all other signals too. The WGN signal was generated and also filtered with the same low pass filter. So, after filtering, it had a bandwidth of one quarter of the sampling rate too. With these preprocessing steps all signals were scaled concerning their bandwidth and their sampling rate, respectively. Additionally, the signal power was scaled. All simulations were performed on a PC with MATLAB.

# 3. Discrimination features based on statistical measures

Several features are considered which are based on measures of the moments $\mu$ and $\sigma$ and/or percentiles $P_y$. The percentile $P_y$ is the resulting abscissa value of a preselected ordinate value $y$ of a distribution function, e.g., for $y = 50\%$ the median $P_{50}$ results. The inclusion of tests for specific distribution functions like chi-square test or Kolmogorov-Smirnov test was abandoned because these tests turned out to be not robust enough for a reliable discrimination of the different signal types. The selected features are: the coefficient of variation *VARCO*, the skewness *SKEW*, the kurtosis *KUR*, and three alternative measures comparable to the first three ones but derived by using several percentile values, *VARCOAL*, *SKEWAL*, and *KURAL*. The first five measures are usual ones for statistical applica-

tions [3] while the last one is an own composition. The six features are:

$$VARCO = \frac{\sigma_\rho}{\mu_\rho} \qquad (1)$$

with $\rho$ – signal magnitude,

$$SKEW = \frac{E\left\{(\rho - \mu_\rho)^3\right\}}{\sigma_\rho^3}, \qquad (2)$$

$$KUR = \frac{E\left\{(\rho - \mu_\rho)^4\right\}}{\sigma_\rho^4} - 3, \qquad (3)$$

$$VARCOAL = \frac{P_{75} - P_{25}}{P_{75} + P_{25}} \qquad (4)$$

with $P_y - y$ percent percentile
of the distribution of $\rho$,

$$SKEWAL = \frac{\mu_\rho - P_{50}}{\sigma_\rho} \qquad (5)$$

with $-1 \leq SKEWAL \leq 1$,

$$KURAL = \frac{1}{6}\frac{(P_{37.5} - P_{25})_{nor}}{(P_{37.5} - P_{25})} + \frac{1}{3}\frac{(P_{50} - P_{37.5})_{nor}}{(P_{50} - P_{37.5})}$$

$$+ \frac{1}{3}\frac{(P_{62.5} - P_{50})_{nor}}{(P_{62.5} - P_{50})} + \frac{1}{6}\frac{(P_{75} - P_{62.5})_{nor}}{(P_{75} - P_{62.5})} - 1. \qquad (6)$$

All features measures are not evaluated for the complex signal values but for their magnitudes $\rho$ because the exact synchronization to the signal was not yet done at this level of signal processing. So, some inaccuracies in the preceding estimation of the centre frequency and its compensation influence the results only marginally. The aim of using these features is the discrimination between strong single channel signals and OFDM, multitone signals or noise like signals. The discrimination between OFDM or multi-tone signals and noise like signals is not possible with these features. This discrimination will be carried out with another feature which will be described in the next section.

In the following, the results of the six features are discussed. In Fig. 2 the results of the coefficient of variation, *VARCO*, are depicted. *VARCO* has small results if the standard variation of the considered variable $\rho$ is small compared to its mean. For the strong single channel QPSK signals, signals 1 and 2 (compare Table 1), *VARCO* is comparatively small. The resulting values increase for the QPSK signals with decreasing *SNRs* (signals 3 and 4) until a value above 0.5 is reached for the magnitude of WGN. The theoretical value of a Rayleigh distributed variable is 0.5227 and is depicted in Fig. 2 with a dashed line. Without regard to the signals 19, 20, and 21 the results for the OFDM signals are all $> 0.45$. So, a decision level for discrimination from single channel signals has to be set to a value between 0.37 and 0.45 depending on the accepted error rate.

The signals 19, 20, and 21 with their comparatively low resulting values belong to those considered HF OFDM types with the higher channel numbers (64 and 93). Apparently, they have a smaller amplitude variance.
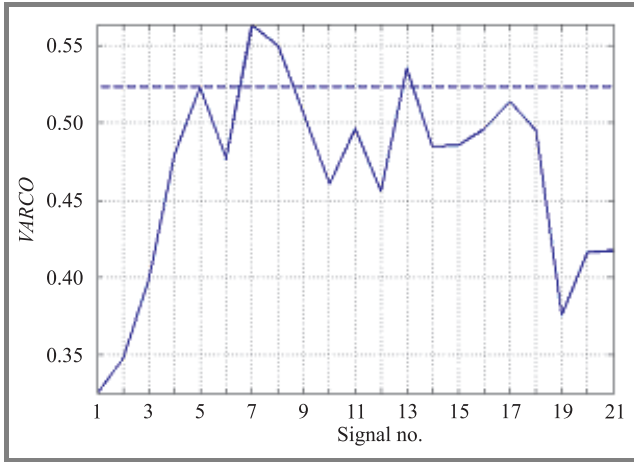


*Fig. 2.* Coefficient of variation.

From a theoretical point of view it may be interesting to have some information about the variance of the *VARCO* values of the particular signals themselves, i.e., their intra signal variance. But, we found out that the statistical variance of a single signal is smaller normally than the variance caused by the different considered signals, i.e., the intra signal variance is smaller than the inter signal variance. So, to keep the clearness of the picture and not to be forced to divide the available real-world signal samples into shorter segments the mean values of the whole signal samples are estimated and depicted only. The same facts are also valid for the other discrimination features which are discussed in the sequel.
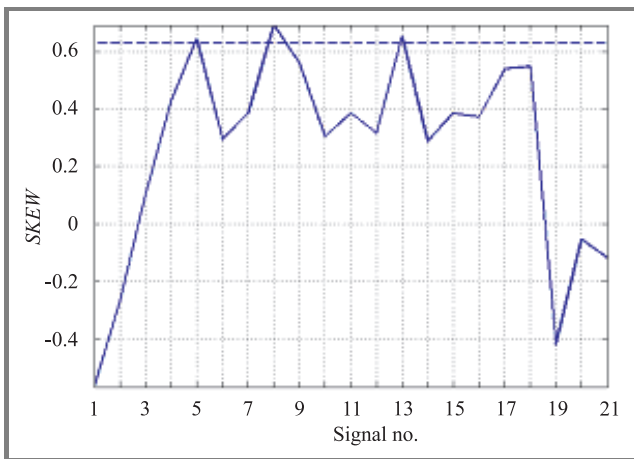


*Fig. 3.* Skewness.

In Fig. 3 the results of the skewness, *SKEW*, are depicted. The skewness is zero for symmetrically distributed variables. It is negative if the distribution density function is skewed to the left and positive if it is skewed to

the right. The Rayleigh distributed variable resulting for the magnitude $\rho$ of a complex WGN has a skewness of 0.6311 which is indicated in the figure by a dashed line and approximated by signal 5. As observed in Fig. 3, the principal arrangement of the results is similar to that for the *VARCO* results. Here, a decision level of about 0.2 seems to be adapted to discriminate most of the considered OFDM signals.
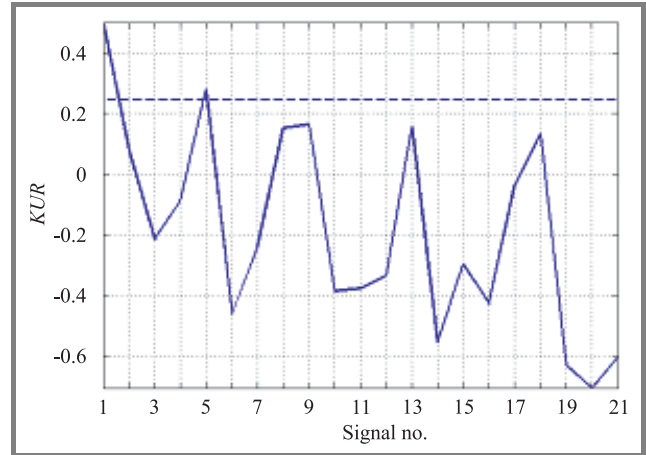


*Fig. 4.* Kurtosis.

Figure 4 shows the results of the kurtosis, *KUR*. The kurtosis is a measure of flatness of a distribution density function near its centre. Positive values are sometimes used to indicate that a density is more peaked around its centre than a normal curve and negative values could indicate that a density is more flat around its centre than a normal curve. The kurtosis of a Rayleigh distributed variable is 0.2451 which is indicated in the figure by a dashed line and approximated by WGN, signal 5. The results in Fig. 4 indicate that the QPSK signal without noise (signal 1) has a more peaked density than the other signals. But, *KUR* has a very large inter signal variance and seems to be not well suited for discrimination. A reason will be that the fourth moment used for its computation is too sensitive to
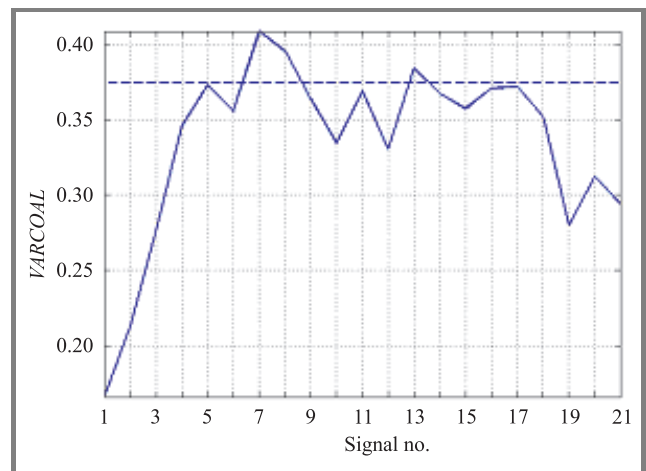


*Fig. 5.* Alternative coefficient of variation.

33

variations of the channel and/or the content of signal information.

The next feature, an alternative coefficient of variation, *VARCOAL*, is computed with the 25% and 75% percentiles and has principally similar results as the coefficient of variation *VARCO* in Fig. 2 but with less inter signal variance of the results, see Fig. 5. So, *VARCOAL* seems to be somewhat better suited as a discrimination feature than the original coefficient of variation *VARCO*. With the dashed line the theoretical result of a Rayleigh distributed variable is depicted again.

The results of an alternative measure of skewness, *SKEWAL*, computed by using not only the mean and the standard deviation but also the median ($P_{50}$) are depicted in Fig. 6. The results are similar to those of the original skewness, i.e., the OFDM signals 19 to 21 cannot be separated.
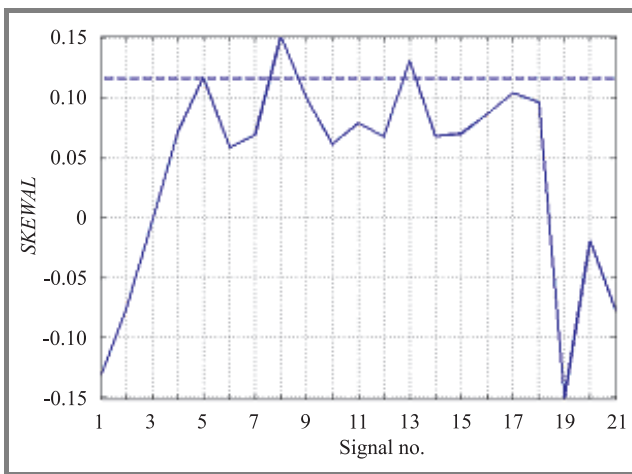


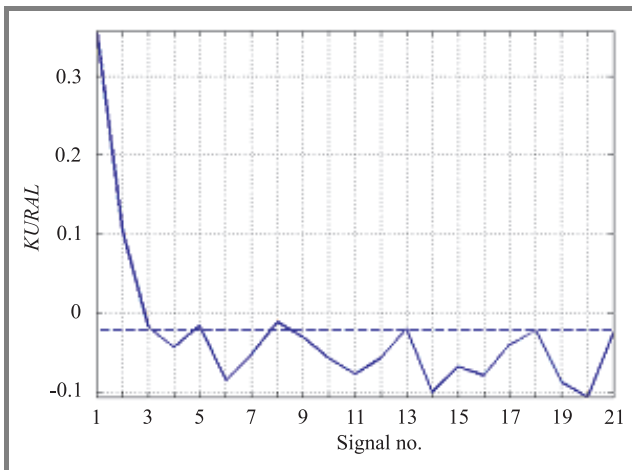***Fig. 6.*** Alternative measure of skewness.



***Fig. 7.*** Alternative measure of kurtosis.

The results of an alternative measure of kurtosis, *KURAL*, composed of percentiles together with an appropriate weighting (see Eq. (6)) are depicted in Fig. 7. The inter signal variance of the OFDM signals and WGN is

comparatively low. This feature seems well suited for discriminating between OFDM respectively WGN and strong single channel signals.

In Figs. 6 and 7 the theoretical results of a Rayleigh distributed variable are again depicted with dashed lines.

## 4. Feature from evaluation of the spectrum shape

The important remaining task is the discrimination between OFDM on one side and a noise like signal, e.g., signal 5 (WGN), or signals without the specific picket-fence shape of their spectra on the other side. An appropriate feature is developed by evaluating the spectral shape. Therefore a spectrum estimate has to be made available. But, this can be taken from the preprocessing procedure where the spectrum has been estimated for the spectral segmentation mentioned in Section 2. The spectrum has to be limited sharply to that part containing high power density. From that spectrum part the cepstrum is computed: the logarithmic spectrum values are transformed with the Fourier transform, i.e., the digital Fourier transform (DFT) in the simulation. The use of the magnitude values of the spectrum makes this computation a non-linear operation. If the analysed spectrum shape has a regular ripple structure, which is typical for many OFDM signals, a significant peak is observed in the cepstrum. The abscissa value of the peak corresponds to the number of periods of the ripple. It is found out that a more distinct peak appearance is reached in the cepstrum in general if the logarithmic spectrum values are weighted with a window function before computing the DFT. For the simulations a Hanning window is used.
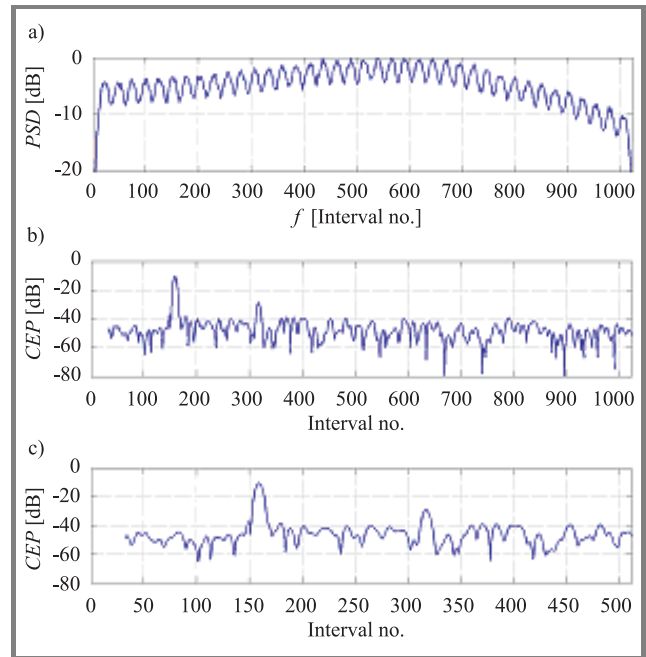


***Fig. 8.*** The HF OFDM signal used for Fig. 1: (a) spectrum; (b) cepstrum, zoom 1; (c) cepstrum, zoom 2.

As an example, the relevant part of the spectrum and the cepstrum of the same HF OFDM signal used for Fig. 1 are shown in Fig. 8. Figure 8a shows the relevant spectrum part, Figs. 8b and 8c depict cepstrum results with different zooms. For a better resolution the cepstrum values are computed with an interpolation factor of four (DFT length of 4096). In the cepstrum graphs a significant peak is observed at the abscissa value of approximately 157. After dividing by the interpolation factor of four the result is 39.25 which is a good estimate of the number of frequency channels of this signal which is 39. The detection and evaluation of such a significant peak can be done automatically. The maximal cepstral peak or, if existing, the two largest peaks with significant level have to be detected ignoring the cepstrum part at the low interval numbers which is not relevant for finding the interesting channel number of an OFDM signal. As a measure of quality of the significant peak its contrast is determined. The contrast is defined here as the difference between peak and maximal side-lobe level (in dB). The maximal side-lobe level is searched within a range of +/– the peak width beside the low ends of the interesting peak.

For comparison purposes the corresponding results of the WGN signal are depicted in Fig. 9. As expected, no significant peaks are observed in the corresponding cepstrum. So, a regular ripple structure in the signal spectrum is not indicated.
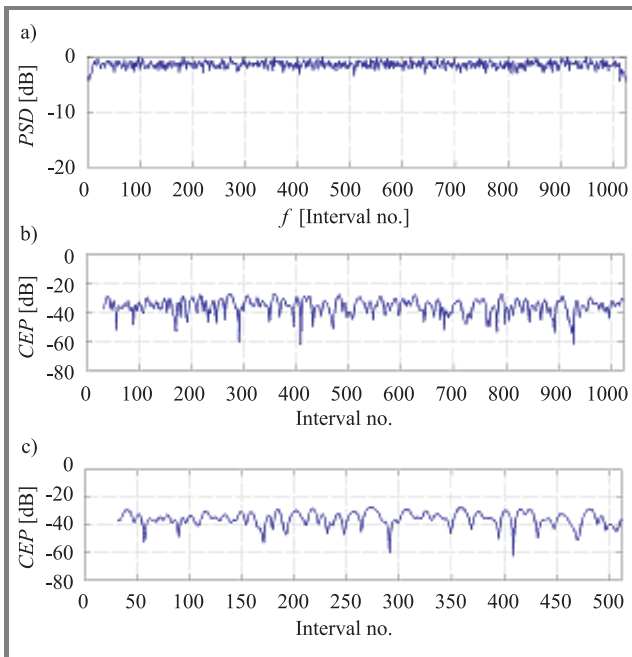


**Fig. 9.** The WGN signal: (a) spectrum; (b) cepstrum, zoom 1; (c) cepstrum, zoom 2.

For all considered signals the appropriate contrast values are depicted in Fig. 10. Signals 1 to 5 do not show remarkable contrast values. Signal 1 to 4 are the QPSK signals and signal 5 is the WGN signal. On the other hand,

the OFDM signals, signals 6 to 21, show significant contrast values although with an appreciable inter signal variance. The signals 11 and 12 with the smallest contrast results are OFDM signals in idle mode, i.e. in non-traffic mode. Frequently, those signal modes do not have well stamped ripple structures in their spectra and the contrast values in their cepstra are less significant. The results of the signals 11, 15, and 21, indicated with bold black star symbols and linearly connected with the other results, are the respective largest contrasts found. But, these contrast values do not correspond to those cepstral peaks which are adjoined to the numbers of system traffic channels these systems have.
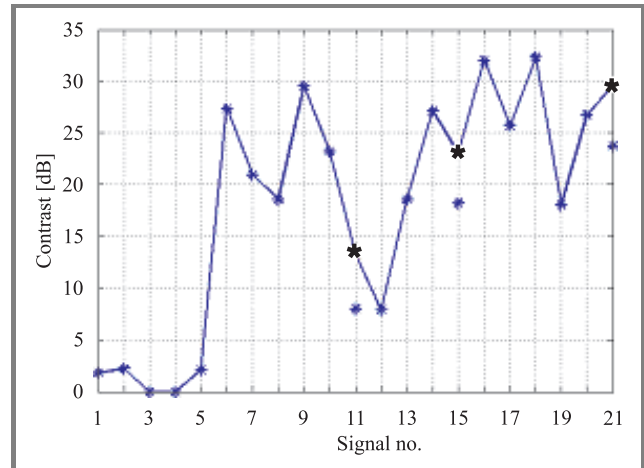


**Fig. 10.** Contrast of significant cepstral peaks.

The reason is that the observed signal samples are some of those ones with idle mode or partly idle mode. Typically, idle mode signals have many lines, sharply peaked, in their spectra. Consequently, the cepstral peaks with the largest contrast values belong to the spectral patterns with the sharply peaked lines. But, for those signals, the cepstral peaks corresponding to the system channel numbers are the ones with the second large values. The results are indicated in Fig. 10 with isolated star symbols. To sum up, the respective maximal contrast values are connected linearly and the contrast values corresponding to the system channel numbers are always depicted with small star symbols.

As observed from Fig. 10, for the discrimination between the considered OFDM signals and other signal types a decision level of 7 to 10 dB for the contrast would be appropriate. By evaluating not only the contrast of the cepstral peaks but their abscissa values too, the channel number and, for idle mode OFDM signals or other signal types, the spectral peak structure can be determined.

With the described feature, not only the discrimination between OFDM and a noise like signal (WGN) is possible but also the discrimination between OFDM and the other signal types which have no significant regular spectral ripples like QPSK. As result, this discrimination feature is an efficient one.

# 5. Conclusions

With the developed discrimination features an automatic recognition of OFDM signals becomes possible. The discrimination capabilities of the considered features are different.

From the two coefficients of variation the alternative one computed with percentile values shows a smaller inter signal variance and, therefore, it is more appropriate for discrimination. The two measures of skewness can discriminate most of the OFDM signals but fail for some types with higher channel numbers. The normal kurtosis measure is less suited for discrimination. Apparently, it is too sensitive to different channel conditions and/or transmitted signal mode (traffic or idle). Contrary to the original kurtosis, the alternative measure composed of percentile values results in a well suited discrimination feature with small inter signal variance. The last feature considered is obtained from evaluation of the spectrum to identify the picket-fence shape which is typical for many OFDM signals. This efficient feature is developed by computation of the cepstrum and evaluation of the largest peaks detected herein. The contrast values of these peaks, exceeding preset decision levels, are used.

Additionally, the number of frequency channels or the structure of the spectrum can be estimated from the abscissa values of the significant cepstral peaks. This feature also discriminates between OFDM and noise like signals.

To develop a complete automatic recognition procedure, the following further steps need to be performed: consideration of more signal samples and tests including synthetically generated signals too, weighting and fusion of the selected discrimination features, and choosing an appropriate classification algorithm.

# References

[1] F. Liedtke, "Adaptive procedure for automatic modulation recognition", *J. Telecommun. Inform. Technol.*, no. 4, pp. 91–97, 2004.

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

[3] A. M. Mood, F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, 3rd ed. New York: McGraw-Hill, 1974.

**Ferdinand Liedtke** received his Ph.D. He works for the Research Institute for Communications, Information Processing, and Ergonomics (FKIE) of the Research Establishment for Applied Science (FGAN) in Germany. His main fields of activity are 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 recognizers offered by national and international companies are based on the concepts worked out at the FGAN-FKIE. He is also interested in modern radio systems and their susceptibilities. A further field of work is the processing of mass data.
e-mail: liedtke@fgan.de
Research Institute for Communications,
Information Processing, and Ergonomics (FKIE)
Research Establishment for Applied Science (FGAN)
Neuenahrer st 20
D-53343 Wachtberg-Werthhoven, Germany

**Ulrike Albers** received her M.Sc. in mathematics. She works for the Research Institute for Communications, Information Processing, and Ergonomics (FKIE) of the FGAN in Germany, on the topics detection and automatic modulation recognition in combination with the development and test of complex algorithms by means of various MATLAB tools. She is an author and co-author of FKIE research reports and conference papers.
e-mail: albers@fgan.de
Research Institute for Communications,
Information Processing, and Ergonomics (FKIE)
Research Establishment for Applied Science (FGAN)
Neuenahrer st 20
D-53343 Wachtberg-Werthhoven, Germany