Overview of frequency bandwidth determination techniques of useful signal in case of leaks detection by correlation method

V A Faerman\textsuperscript{1a}, V S Avramchuk\textsuperscript{1b}, E E Luneva\textsuperscript{1c}

\textsuperscript{1} Tomsk Polytechnic University, Tomsk, Russia
\textsuperscript{a} e-mail: fcsevladfaerman@gmail.com
\textsuperscript{b} e-mail: avs@tpu.ru
\textsuperscript{c} e-mail: lee@tpu.ru

Abstract. In this paper an overview of useful signal detection methods on the background of intense noise and limits determination methods of useful signal is presented. The following features are considered: peculiarities of usage of correlation analysis, cross-amplitude spectrum, coherence function, cross-phase spectrum, time-frequency correlation function in case of frequency limits determination as well as leaks detection in pipelines. The possibility of using time-frequency correlation function for solving above named issues is described. Time-frequency correlation function provides information about the signals correlation for each of the investigated frequency bands. Data about location of peaks on the surface plot of a time-frequency correlation function allows making an assumption about the spectral composition of useful signal and its frequency boundaries.

1. Introduction

Recovery of useful signals on a background of random noise is one of the most actual problems of digital signal processing (DSP). Despite the fact that so far, many methods and techniques were developed in order to solve this problem, they all have significant limitations in use and none of them is universal. The problem of useful signals detection on the background of random noise was widely considered in the middle of last century due to the rapid development of radio and radar systems [1]. Various method of mathematical statistics were the foundation of techniques aimed to solve this problem, and field of radio engineering related to the development of these techniques was called the theory of optimal signal reception [1]. Currently, the issue of signal detection with a priori unknown format is typical problem for radiolocation branch as well as for non-destructive testing [2, 3, 4]. Specifically, one of the problems of signal detection and determination of its unknown parameters is the problem of detecting leaks solved by the correlation method [2].

2. Setting of signal detection problem and estimating signal’s parameters problem

In a generalized sense the problem of optimal signal reception can be formulated as follows: on the supposition of in a priori known some characteristics of the useful signal, noise, as well as their functional interaction, it is necessary to obtain optimal receiving or decision-making device. This device should in the best way reproduce a signal or make decision with the least errors [1]. Depending on the amount of a priori available data, several tasks of varying difficulty receiving signals were distinguished [1]:

- simple detection of the useful signal that is completely known;
- complex signal detection with unknown parameters without measuring these parameters;
- signal detection with measuring its parameters.

Traditionally, much attention was directed to the last group of tasks due its complexity, as well as practical significance for radiolocation branch.

A common radiolocation task of signal detection with measuring its unknown parameters is a task of correlation signal \( s(t, \lambda) \) reception, having one unknown parameter \( \lambda \). It is assumed that a receiver device takes up a signal \( \tilde{\zeta}(t) \), and this signal is a mixture of a detectable signal \( s(t, \lambda) \) and additive...
random noise \( n(t) \). Below the case of discrete signal processing is considered. So signal \( \xi(t) \) and \( s(t, \lambda) \) are represented by time series containing the samples that are equally spaced in time by a value \( \Delta \):

\[
i = \xi(t_i), \quad s_i(\lambda_j) = \xi(t_i) \quad (t_i = i \cdot \Delta, i = 1, 2, ..., N - 1).
\]

Total number of available sample units is determined by the observation time \( T \) and the discretization interval \( \Delta \) according to the following formula:

\[
N = \frac{T}{\Delta}.
\]

Methodics of solution described problem comes down to obtaining and analyzing the correlation function \([1]\).

\[
R_{\xi s}(j) = \sum_{i=0}^{N-1} \xi_i(j) = \sum_{i=0}^{N-1} \xi(t_i)s(t_i, \lambda_j),
\]

where \( s(t_i, \lambda_j) \) — reference implementation of useful signal when \( (\lambda = \lambda_j) \); \( r_{\xi s}(j) \) — correlation function of mixture of signals \( \xi(t_i) \) and reference implementation of useful signal \( s(t, \lambda_j) \). Further analysis of correlation function is reduced to obtaining its maximum, that, in simple case, corresponds to the value of requiered parameter \( \lambda \). Usually under requiered parameters serves: time delay of arrival reflected signal \( \tau_0 \), which is used in order to determine the distance to the object location; Doppler frequency \( f_0 \), which is used in order to determine the velocity of the object.

Despite the fact that Eq. 1 can be simplified in case of the signal \( s(t, \lambda_1, ..., \lambda_n) \) detection, that depends on many unknown parameters \( \lambda_1, \lambda_2, ..., \lambda_n \):

\[
R_{\xi s}(j, k, ..., m) = \sum_{i=0}^{N-1} \xi(t_i)s(t_i, \lambda_1, \lambda_2, ..., \lambda_m),
\]

the complicity of proceed operations is rising significantly and this leads to the need to increase number of calculation devices (correlators) as well as to make the receiver device more complicated \([1]\).

3. Correlation method of leak location detecting in a pipeline

Description of the method. Schematically location leaks detection by correlation method is represented in Fig. 1. Since the sensors and signal source are lain out in single line, the location of signal source is uniquely determined by the values of distance from both sensors. The propagation velocity of the signal is determined dependant on solving task, and so usually it can be measured \([5]\). Thus, the problem reduces to determining the delay signal on sensor by correlation method.

![Figure 1 – Scheme of location leaks detection by correlation method](image)

\( r_1, r_2 \) — distance from the signal source to the sensor; \( d \) — distance between sensors

The main feature of the correlation method lies in simultaneous use of a pair of sensors. Suppose the sensors, during the interval \( T \), fix signals that are arrived from equally separated moments of time \( \Delta \). Then in any discrete time \( t_i \) \( (i = 0, 1, ..., N - 1, \) where \( N \) — whole number of sample units \) a signal that registered by the first sensor can be represented as following:

\[
n_1(i) = s_0(i - j_1) + n_1(i),
\]

where \( n_1(i) \) — noise component of signal from the first sensor; \( j_1 = \tau_1/\Delta \) — signal lag from the first sensor \( \tau_1 \) that is measured in sample units. In a similar way, for second sensor signal:

\[
n_2(i) = s_0(i - j_2) + n_2(i).
\]
Correlation function should be calculated as following [3]:
\[ r_{12}(j) = F^{-1}(F[x_1(i)]F^*[x_2(i)]), \]  \hspace{1cm} (4)
where \( F \) - direct discrete Fourier transform (DFT); \( F^{-1} \) – inverse DFT; \( F^* \) - complex conjugate representation of the direct DFT results. Suppose that noise components of signal from sensors \( n_1 \) and \( n_2 \) are not correlate that allow representing Eq. 4 as a following sum:
\[ r_{12}(j) = r_{n0n0}(j) + r_{n1n0}(j) + r_{n2n0}(j) + r_{n1n2}(j). \]
It is known that the following maximum value \( \max (r_{n0n0}(j)) \) corresponds to \( r_{n0n0}(j_2 - j_1) \) [2, 5] in other words to the difference in arrival time of signals to sensors.

**Limitation of the method.** In order to detect a maximum of the correlation function, and this maximum is informative, the following condition has to be fulfilled:
\[ r_{n0n0}(j_2 - j_1) \gg \sqrt{D[r_{n1n0}(j) + r_{n2n0}(j) + r_{n1n2}(j)]}, \]
where \( D \) – variance of random variable. Thus, an unknown parameter is defined by the ratio of peak determinate component value of correlation function \( (r_{n0n0}(j)) \) and root mean square (RMS) value sum of its random components \( (r_{n1n0}(j) + r_{n2n0}(j) + r_{n1n2}(j)). \)

**Techniques that reduce RMS value of random components.** In practice there are different techniques of RMS value reducing of signal random components. The obvious technique is consist in increasing sample size \( (N) \), that allow weakening random components owing to averaging properties of the correlation sum. However, increasing the sample size in some cases is impossible because of fast Furrier transform limitation on number of sample units, which is used as the fastest way of correlation function calculation [2].

Another simple and effective technique to reduce the RMS value of signal random components at the output of a correlator is a digital filtering (usually band or band-rejection filtering) at the input of the correlator. But, this technique implementation has some difficulties, since signal spectrum usually is unknown [6]. Therefore diverse techniques of identifying bandwidth of spectral localization of signal are used for following validated filter settings.

### 3. Frequency bandwidth determination techniques of useful signal.

The simplest frequency bandwidth determination techniques consist in obtaining cross-amplitude spectrum [4]. Cross-amplitude spectrum \( S'_{12}(k) \) can be obtained by multiplying frequency samples and amplitude signal spectrums of sensors \( |X_1(k)|, |X_2(k)| \):
\[ S'_{12}(k) = |X_1(k)| \cdot |X_2(k)| \]  \hspace{1cm} (5)
where \( k \) – number of spectrum sample \((k = 0, 1, ..., N/2 + 1)\). In practice instead of Eq. 5 is widely used the following equation:
\[ S'_{12}(k) = |F[x_1(i)]F^*[x_2(i)]|. \]  \hspace{1cm} (6)

The advantage of Eq. 6 is the use of intermediate data that obtained in the correlation function calculation according to Eq. 4. Noticeable peaks as well as , as well as high values of cross- amplitude spectrum are signs of the presence of the useful signal in the considered frequency band. Normally, cross-amplitude spectrum is more informative than the amplitude spectra of signals from each of the sensors. In the Fig. 2 amplitude spectra of signals from sensors and cross-amplitude spectrum are represented. Test signal is the sum of normally distributed random quasi-white noise and narrowband signal with frequency range 10…12 kHz.

The disadvantage of this technique is its sensitivity to the spectrum of the noise component at the input of the correlator device. If the noise energy is not uniformly distributed, but concentrated in some parts of the spectrum, it cannot be excluded the wrong choice of frequency bands for further analysis.
Figure 2 – Amplitude spectra of signals from sensors $|X_1(f)|$, $|X_2(f)|$ and their cross-amplitude spectrum $S'_{12}(f)$

4. Coherence function analysis

Another common way to obtain a frequency bandwidth of useful signal is based on calculating coherence function. The coherence function is usually defined as the ratio of signals cross spectrum and root of the product of their own spectrums:

$$\gamma(k) = \frac{|E(S_{12}(k))|}{\sqrt{E(X_1^2(k))E(X_2^2(k))}}, \quad (7)$$

where $E$ – averaging operator; $X_1^2(k), X_2^2(k)$ – eigen spectrum of signals $\xi_1(i)$ and $\xi_2(i)$ properly; $S_{12}(k)$ – cross spectrum of signals ($S_{12}(k) = F[\xi_1(i)]F^*[\xi_2(i)]$). Also in practical spectral analysis is often the square of the coherence function calculated. When coherence function is calculated according to Eq. 7 or square coherence function, the value of signal spectrum should be averaged by some sample. If the sample is represented by a single set of spectrum, Eq. 7 takes the value 1.

Therefore, in practice of coherent analysis, the square of the coherence function is calculated by the following equation:

$$\gamma^2(k) = \frac{|\sum_{q=0}^{Q-1} S_{12}(k)|^2}{\sum_{q=0}^{Q-1} X_1^2(k) \sum_{q=0}^{Q-1} X_2^2(k)}, \quad (8)$$

where $Q$ – number of spectrum sets of signals, applicable in coherence function calculating. Each set of spectra requires $N$ samples of the measured signal, thus total sample size required coherence function calculation is $Q$. 

In the fig. 3 coherence function plot for different value of Q is represented. Test signal is the sum of normally distributed random quasi-white noise and narrowband signal with frequency range 14…17.25 kHz. The coherence function in Fig. 3a contains no useful information. On the contrary, a region of spectrum localization of useful signal, as shown in Fig. 3b, is well defined.

Figure 3 – Coherence function for different number set of short-range spectrum of signals Q:

a) $Q = 5$; b) $Q = 20$.

Thus, coherence function is an analogue of the correlation function, in case of representing the correlation of processes in frequency domain [7]. In addition, the region of high values of a coherence function indicates the presence of related and not random signals. This fact can be used in problems of signal detection [6].

There are following disadvantages of analysis approach based on coherence function: requirement in the accumulation of additional data about signals; possible occurrence of false high values of the function in case of some problems determined by the existence of several vibration modes [8].

5. Time-frequency correlation analysis

Another way of correlation analysis that allows extracting and visualizing information about spectral composition of a signal is the use of time-frequency correlation functions [6]. Unlike with common functions, described above, time-frequency correlation functions have the additional argument – frequency. Thus, time-frequency correlation function represents a correlation of signals for each studied frequency band. The spectral composition of a useful signal is estimated by the location of peaks on the time-frequency correlation function plot [9]. For example, time-frequency correlation function plot was obtained for test signal that contains white noise and narrowband useful signal. As it shown in Fig. 4, the useful signal was located in frequency band $4…5.2$ kHz.

The peaks of time-frequency correlation function become more noticeable at frequency bandwidth of useful signal. This can be explained by the fact that the algorithm for obtaining a time-frequency correlation function, represented [6], is equivalent to the simultaneous use of a number of correlators with band-pass basic frequency filter at the input.
6. Summary

Due to the fact that the major correlation method limitation in case of signal detection with a priori unknown format is impossibility to manage with high intensity noise at the input of the correlator. So the most important step of solving this problem is filtering of an input signal.

In order to obtain validate filter selection, spectral bandwidth of useful signal should be localized. In this case the same approaches to perform this step, give different efficiencies depending on a field of application (i.e. leak detection of oil pipelines, water pipelines, etc.). Therefore, the best result can be achieved by combined use of approaches that were described in this work.

References

[1] Tikhonov V I 1983 Optimal Signal Reception, Radio i Svyaz, Moscow.
[2] Ifeachor E C, Jervis B W 2011 Digital Signal Processing: practical approach, John Willey & Sons Ltd., West Sussex.
[3] Beck S B, Curren M D, Stanway R 2005 Pipeline Network Features and Leak Detection by Cross-Correletion Analysis of Reflected Waves, J. Hydraulic Engineering, 131, 715-723.
[4] Bendat J S, Piersol A G 2013 Engineering Applications of Correlation and Spectral Analysis, Wiley-Interscience.
[5] Huandi O, Chue W, Wang A, Guan W 2000 Detecting leaks in plastic pipes, J. AWWA, 92, 82-94.
[6] Avramchuk V S, Goncharov V I, Chan Viet T 2009 Russian Patent 2405163.
[7] Matz G, Hlawatsch F 2000 Time-frequency coherence analysis of nonstationary random processes, proc. of 10th IEEE-SP Workshop on Statistical Signal and Array Processing, Pocono Manor, 554-558.
[8] Ovchinnikov A L, Lapshin B M 2008 Experience of using a leak detector TAK-2005 for service of the municipal pipelines, Bulletin of Tomsk Polytechnic University, 312, 2, 196-202.
[9] Avramchuk V S, Goncharov V I 2013 Time-frequency Correlation Method for Improving the Accuracy in Detecting Leaks in Pipelines, Advanced Materials Research, Trans Tech Publications, Switzerland, 650, 443–446.