Errors Detection Algorithm Description

There are many faults that can arise in an audio transmission channel between the sound source and the listener. The automatic computer surveillance system is able to recognize and report most of the typical errors that can occur in audio broadcast or streaming. The system works in real time without any knowledge of how the sound signal is processed before transmission or streaming. Thus the only way how to detect some types of error in an audio transmission is comparing the output signal of the sound source (reference signal) with the processed and transmitted audio signal received by receiver (transmission signal). In this way the errors detection algorithms can be divided into two groups:

- Single signal error detection – silence, interrupt, overload, hum noise, click noise and clipping.
- Cross signals error detection – delay (synchronization), audio quality comparing, and silence.

The first group of error detection algorithms processes all received signals. After that signals are down-mixed (if it is required) and buffered for synchronization. Synchronized signals are processed by second group of error detection algorithms.

The surveillance system consists of two stages. In the first stage (*init stage*) are signals synchronized and all error detections are disabled except silence detection. After signal synchronization the second stage (*process stage*) is started and all error detections are enabled (except the error detections disabled by user in config file). When synchronization error in any channels is reported, all other error detections of second stage in these channels are disabled. When silence in any channel is reported all error detections are disabled in this channel.

The error detection algorithms are driven by user's settings of parameters in config file. Some kinds of parameters are same for all error detections and some of them are different for particular detection. The same kinds of parameters are described below and the others will be described with appropriate error detection algorithm.

- *Enable* – enables (1) / disables (0) appropriate error detection.
- *Notice only* – defines whether the error is reported as notice (1) or warning (0).
- *Error Number* – number of detected errors before the error is reported (used for multiple error detections, recommended for click noise and clipping detections)
- *Time Window* [*s*] – time window in seconds for errors to raise warning/notice. For example if it is set to 1 and the *Error Number* is set to 3, the system reports notice/warning only when this error occurs more than 3 times within 1 second interval.

Each error detection is reported as warning or notice only (according user settings in config file), the only exception is synchronization error detection which is always reported as warning, and silence detection which has specific behavior described below.

Silence

Silence detection algorithm is based on measurement of input signal level by VU meter. The VU meter follows the ASA C16.5-1954 specification. This standard specifies the meter time behavior (meter ballistics) in this way: the rise time when the level meter reaches a value of 0 dB must be 300 ms when the nominal signal is fed to the meter input and the indicator must fall down during 300 ms again when the nominal input signal is disconnected. When the VU meter level of the signal is lower than the *Threshold* for a time *T* longer than the *Period*, a notice is reported. When silence is detected in all channels of the logical channel group simultaneously, a warning is reported (not in the simple comparing mode). When silence is detected in particular channel only either of reference line or of transmission line, a warning is reported. The figures below illustrate the silence detection.

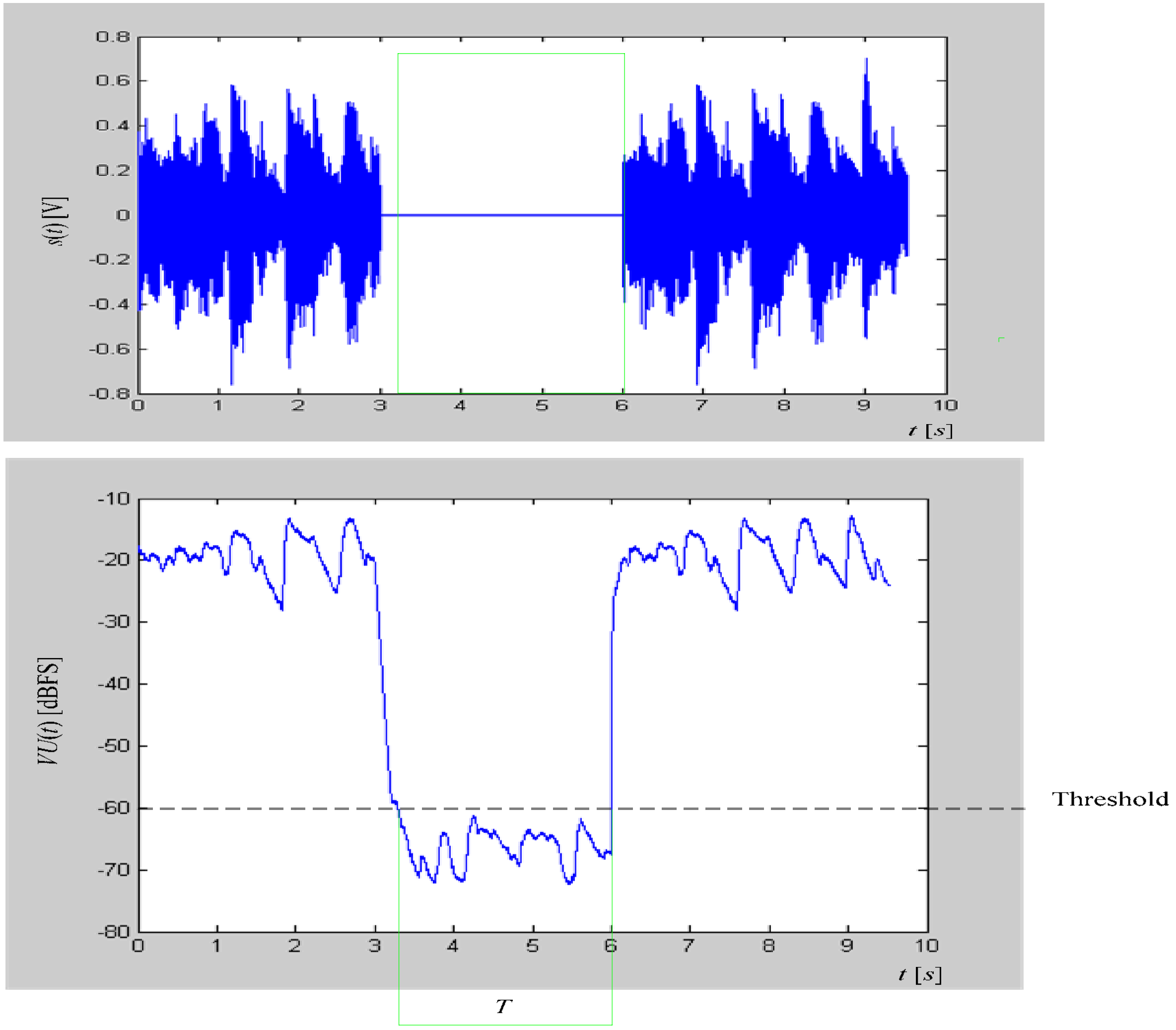

Fig. Signal waveform and its level measured by VU meter. Detected silence is marked by green rectangle.

Parameters:

- *Threshold* [dBFS] - minimal level of the VU meter measurement for non-silenced signal. Default value is -60 dBFS.
- *Period* [s] – maximal time interval for which the signal level can be below the *Threshold* and error is not reported yet. Default value is 5 s.

Interrupt

Interrupt detection algorithm is based on measurement of input signal level by volume unit VU meter same as the silence detection algorithm is. When the VU meter level of the signal is lower than the *Threshold* for a time *T* longer than the *Period*, a notice/warning is reported. The *Threshold* value is lower and the *Period* is shorter than as they are in the silence detection algorithm. The figures below illustrate the interrupt detection.

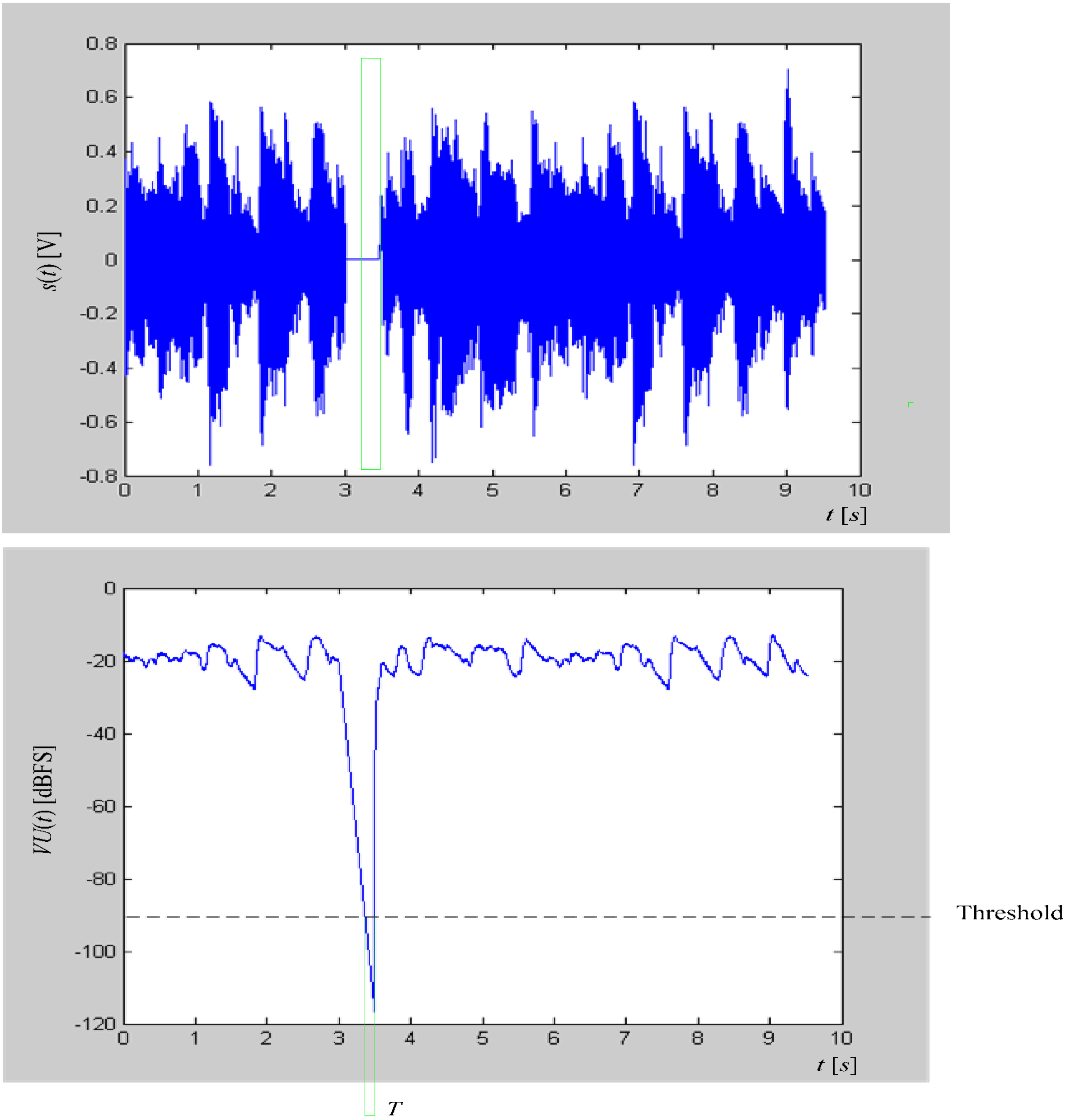

Fig. Signal waveform and its level measured by VU meter. Detected audio interruption is marked by green rectangle.

Parameters:

- *Threshold* [dBFS] - minimal level of the VU meter measurement for non-interrupted signal. Default value is -90 dBFS.
- *Period* [s] – maximal time interval for which the signal level can be below the *Threshold* and the error is not reported yet. Default value is 0.15 s.

Overload

Overload detection algorithm is based on measurement of input signal level by peak program meter PPM. The PPM follows the DIN 45406 standard. This standard defines that the meter must show a value of –1 dB in a period of 10 ms after the nominal signal has been fed to the meter input and it has to fall down at a rate of 20 dB per 1.5 s. When the PPM level of the signal is higher than the *Threshold* for a time *T* longer than the *Period*, a notice/warning is reported. The figures below illustrate the overload detection.

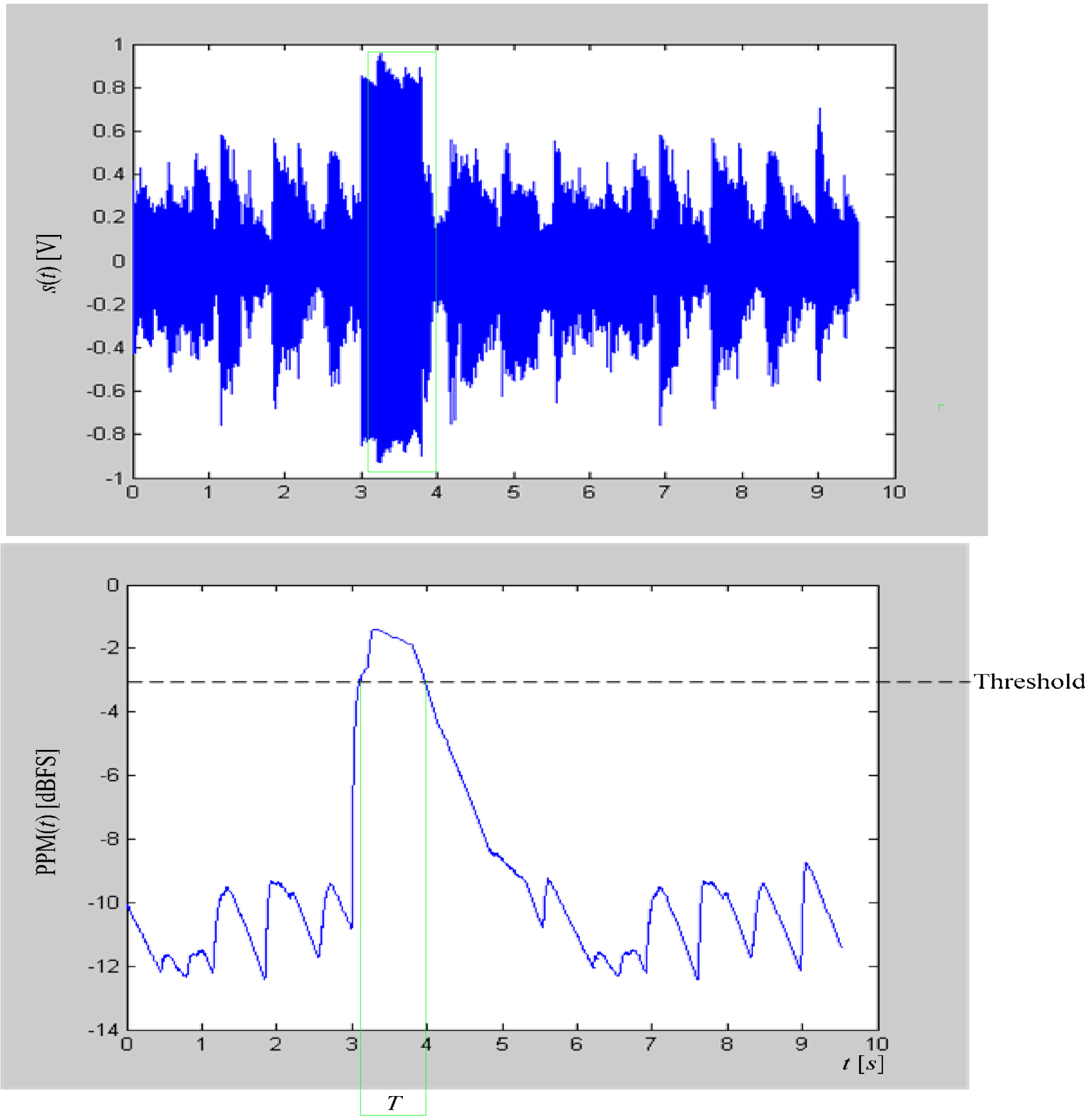

Fig. Signal waveform and its peak level measured by PPM. Detected audio interruption is marked by green rectangle.

Parameters:

- *Threshold* [dBFS] - maximal level of the PPM measurement for non-overloaded signal. Default value is -3 dBFS.
- *Period* [s] – maximal time interval for which the signal peak level can be above the *Threshold* and the error is not reported yet. Default value is 0.2 s.

Hum noise

When the harmonic signal with a frequency of 50/60 Hz and its second and third harmonics (100/120 Hz, 150/180 Hz) with the overall root mean square RMS level above the user-defined level of the signal RMS are recognized in the input signal, a notice/warning is reported. The base frequency 50/60 Hz corresponds to a power supply frequency and can be set in general settings in the config file.

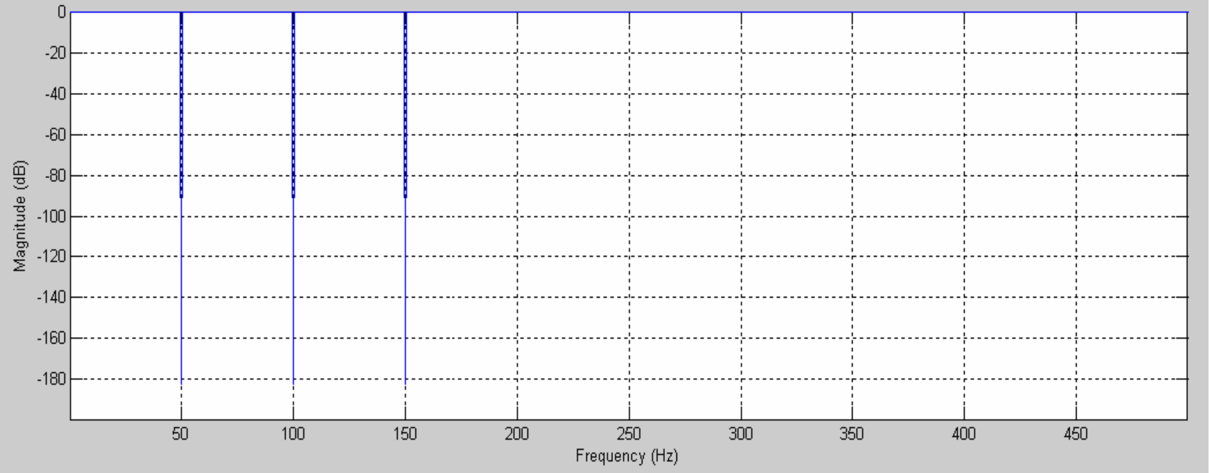

Fig. The frequency response of three notch filter cascade with base frequency 50 Hz.

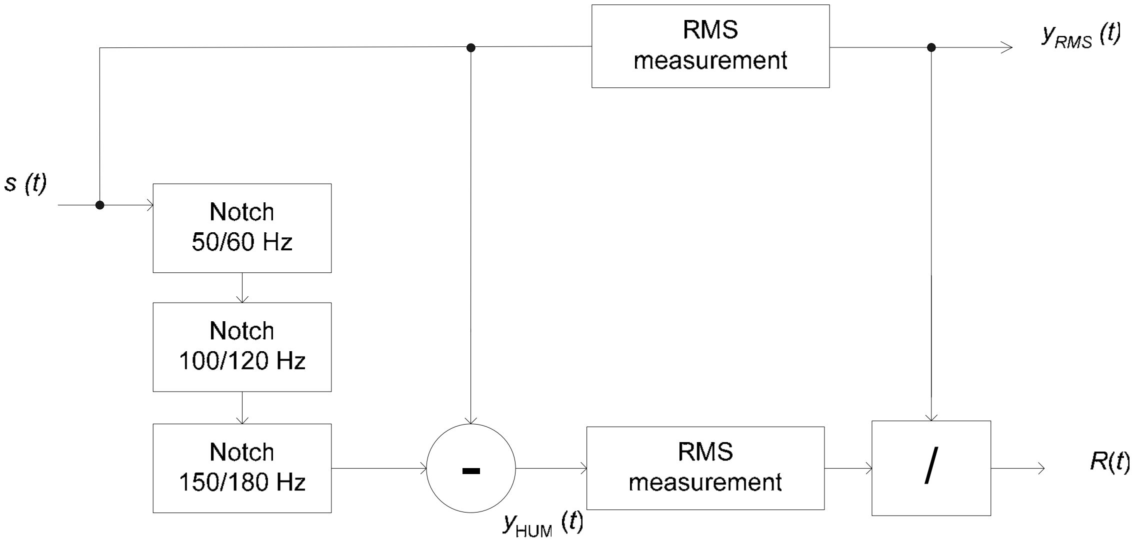

Fig. Hum noise measurement algorithm.

The algorithm uses a cascade of three notch filters with frequency response shown on figure above. The output signal of notch filter cascade is subtracted from the input signal and fed to the input of RMS measurement system and its current output signal value is compared with the current RMS level of the analyzed signal. When the comparison level $R(t)$ of the signal is higher than the *Threshold* for a time T longer than the *Period*, a notice/warning is reported. The figures below illustrate the hum noise detection.

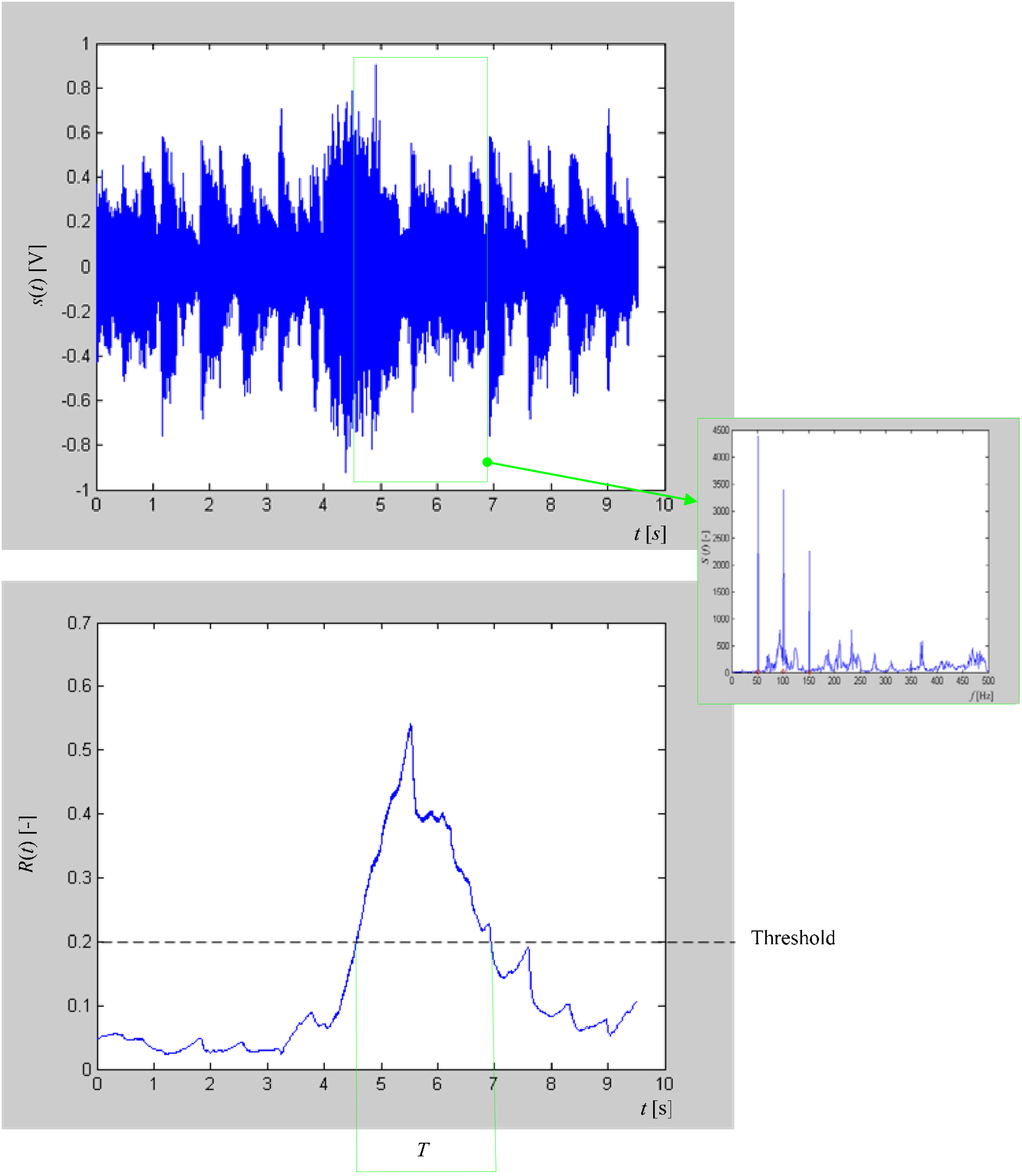

Fig. Signal waveform and comparison of signal RMS level and hum RMS level $R(t)$. Detected hum noise occurrence is marked by green rectangle. In the green rectangle on the left side is shown spectrum of signal with hum noise.

Parameters:

- *Threshold* [-] – maximal ratio of signal RMS and of hum noise RMS for signal without hum noise. Default value is 0.2.
- *Period* [s] – maximal time interval for which the comparison level $R(t)$of signal RMS and hum RMS can be above the *Threshold* and the error is not reported yet. Default value is 1 s.

Click noise

Click noise is expressed by series of so called *clicks* in short time interval. The single click definition can be described as disturbance of a signal time behavior by unexpected step change of particular sample value (in negative or positive direction) and recovering to supposed value in interval of few samples. The click noise detection algorithm is based on the comparison of the signal first derivative values and of their peak function. The click is detected, when the comparison level $R(t)$ of the signal gets over the *Threshold* (in positive or negative direction) and during C samples it gets over *Threshold* with opposite polarity, while C is less than *Count.* For click noise detection is recommended to use multiple click detection in a short time interval (approx. 5-6 click detections during 1 s) to notice/warning to be reported. The figures below illustrate the click noise detection.

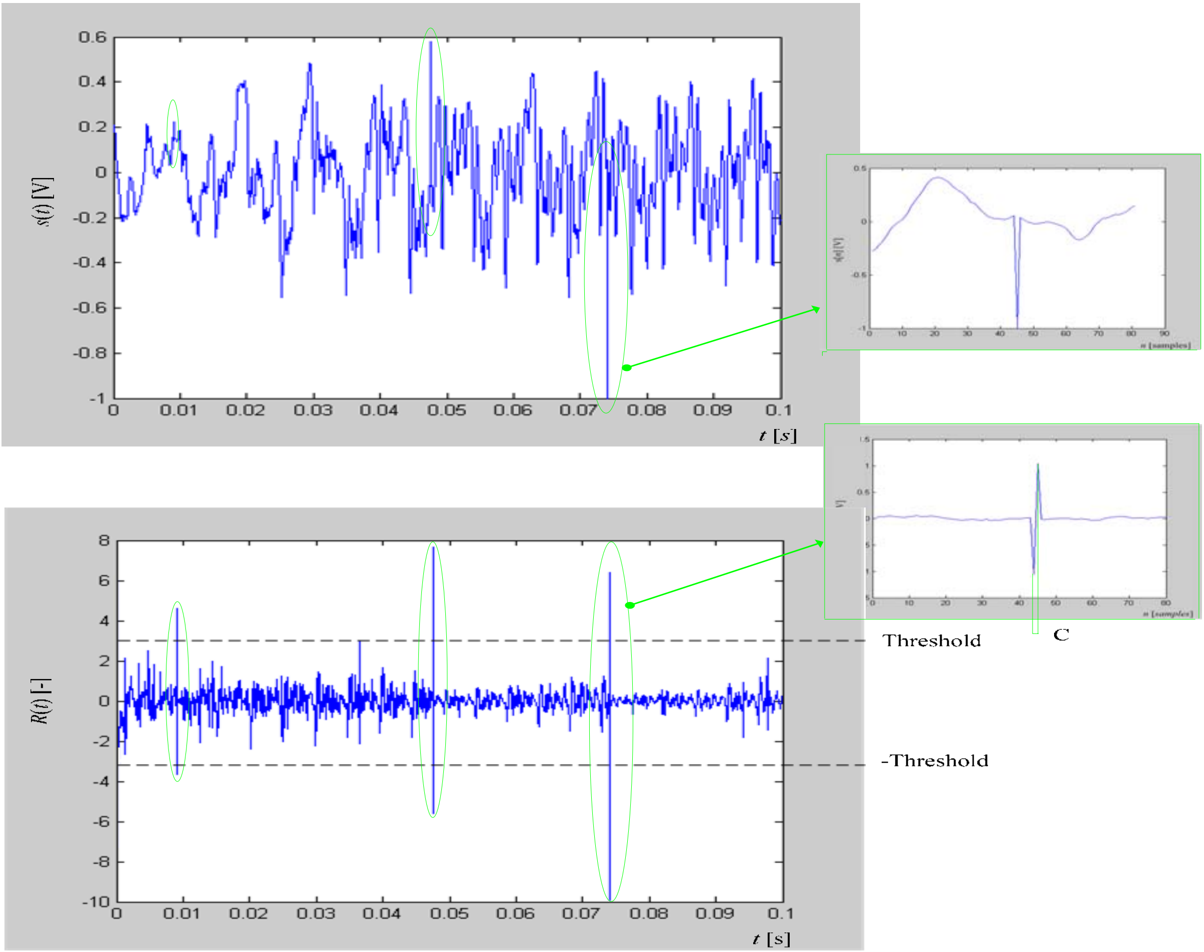

Fig. Signal waveform and comparison of signal first derivative values and of their peak function $R(t)$. Detected clicks occurrences are marked by green ellipsis. In the fist green rectangle on the left side is shown detailed time behavior of signal with click, and in the second green rectangle is shown the first derivate of this signal, where C counts number of samples between two highest peaks of opposite polarity determining beginning of click and recovery from click.

Parameters:

- *Threshold* [-] – maximal positive and minimal negative ratio of difference signal and of its peak function for signal without click noise. Default value is 3.
- *Count* [samples] – maximal number of samples for signal recovery from a click to be reported as click. Default value is 10 samples.

Clipping

The clip detection performs the counting of the occurrence of samples with full-bit digit values. Due to possible attenuation and low-energy noise adding in transmission environment is full-bit digit value replaced by user defined *Threshold* in dBFS, which should be just a little lower than full-bit digital value 0 dBFS. When the number of sequencing samples *C* rises above a user-defined *Count*, a clip is detected. For clipping detection is recommended to use multiple clip detection in a short time interval (approx. 7-68 clip detections during 1 s) to notice/warning to be reported. The figures below illustrate the clipping detection (for better illustration there is used waveform instead of level measurement in dBFS).

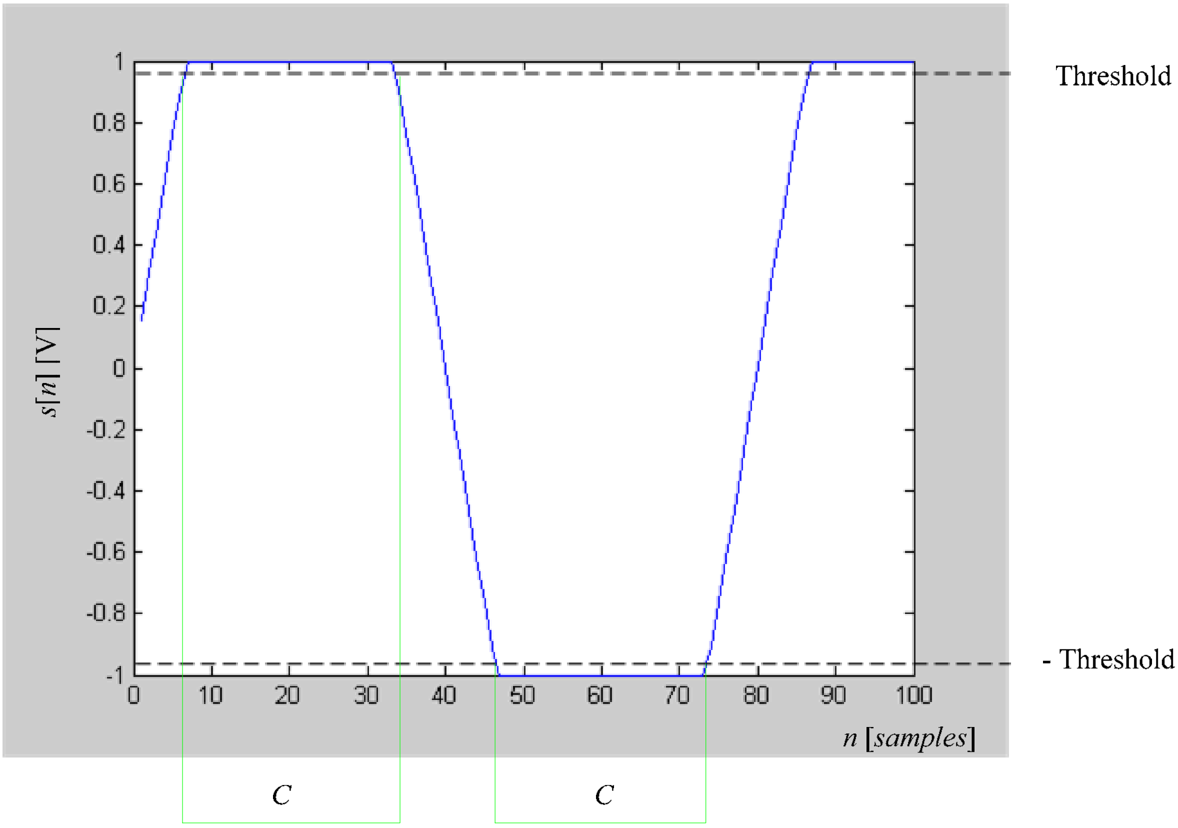

Fig. Signal waveform with two clips. Detected clips are marked by use of green lines.

Parameters:

- *Threshold* [dBFS] – maximal signal level for signal without clipping. Default value is -0.25 dBFS.
- *Count* [samples] – maximal number of sequencing samples for which the signal level can be above the *Threshold* and the error is not reported yet. Default value is 50 samples.

Delay

The fast cross-correlation analysis using the fast Fourier transform (FFT) is used for delay detection. Because of the computing-power demands, only the first channels of the reference and the transmission lines are correlated even if a stereo signal is present at the output of the channel switching block. The algorithm works in two phases: during the initialization phase, long data buffers (approx. 10 s for sampling frequency 48000 Hz) are used for the cross-correlation analysis and the major delay proper is detected. Maximal allowable delay is approx. 5 s for sampling frequency 48000 Hz. The init phase longs approx. 50-60 s, if the delay is found the process phase starts, if not then *init_fail* message is reported and init phase starts again while the delay is not found.

In the process phase signals are synchronized (both channels for stereo signal) and delay changes are tracked. The signal period for the correlation in this phase is about 1.5 s for sampling frequency 48000 Hz. Neighboring segments are overlapped with ¾ of the their length. This eliminates delay detection errors caused by rhythmical musical signals and gives fast results. Short-time averaging of the detected delay gives stable results. Maximal allowable delay change is 100 ms per 1 s in both directions. If the delay change is greater or signals become unsynchronized for more than *Count* segments, synchronization is corrupted and *unmatched* message is reported.

In simple comparing mode is allowable delay defined by *Threshold*, when it is higher synchronization is corrupted and *unmatched* message is reported.

Parameters:

- *Threshold* [s] – maximal allowable delay between reference line and transmition line in simple comparing mode. Default value is 0.05 s.
- *Count* [segments] – maximal number of segments for which it is not possible find appropriate delay and the error is not reported yet. Default value is 3 segments.

Audio quality comparing

Audio quality comparing is measured by use of normalized spectral coefficients cross correlation function *Ncc*(t) between reference and transmission line channels. This function uses correlation of two segments of synchronized signals with length corresponding to the surveillance system input buffers length. When the level of *Ncc*(t) is lower than the *Threshold* for a time *T* longer than the *Period*, a notice/warning is reported. In the simple comparing mode is *Threshold* fixed set to 1.The figures below illustrate the audio quality comparing error detection.

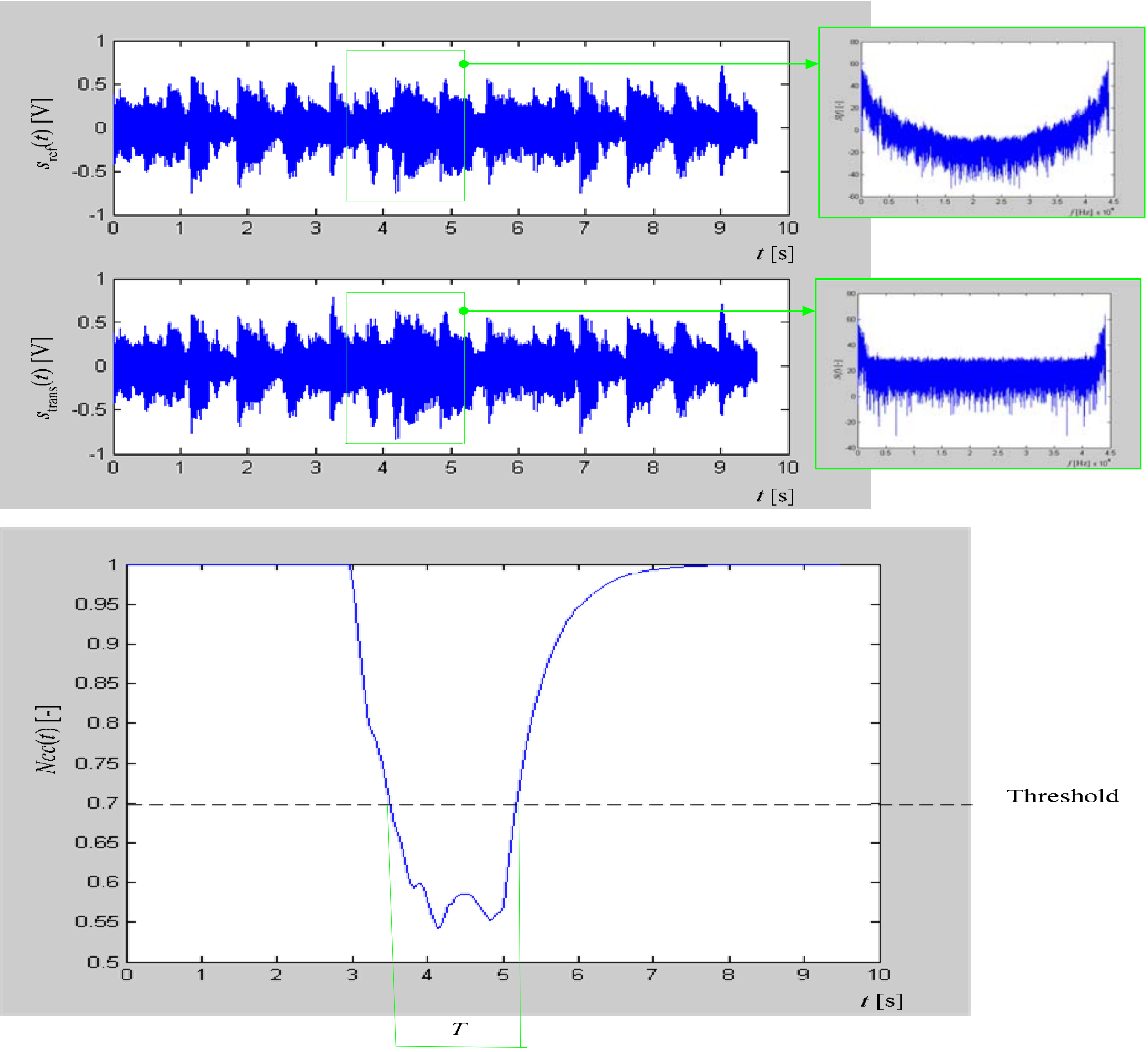

Fig. Waveforms of reference and transmission signal and normalized spectral coefficients cross correlation function *Ncc*(t) between these signals. Detected audio quality loss occurrence is marked by green rectangle. In the green rectangles on the left side is shown spectrum of compared signals where the error is detected.

Parameters:

- *Threshold* [-] – minimal level of normalized spectral coefficients cross correlation function *Ncc*(t) between reference and transmission line channels for non audio quality error signal. Default value is 0.7.
- *Period* [s] – maximal time interval for which the *Ncc*(t) level can be below the *Threshold* and the error is not reported yet. Default value is 1 s.