

# Evaluation of Nonlinear Digital Audio Systems Through Transparency Reduction Using the Combined Test Signal

UDK 621.391.8:004:534-7  
IFAC 5.8.3

Original scientific paper

The well known and established methods used for distortion measurements cannot be used for detecting the errors that appear in digital audio systems based on perceptual coding and decoding. A test stimulus has to be used that has the time-varying properties similar to real signals such as speech and music. Furthermore, the procedure of quality evaluation has to include an estimate of the temporal behaviour of human hearing. For the purpose of evaluation of digital audio systems based on perceptual coding, a new combined test signal was developed with the cumulative distribution of amplitude following the Gaussian distribution. Input and output signals are transformed from the objective domain to the excitation domain. Transparency reduction is defined as a measure of distortion, being the ratio of error excitation and stimulus excitation.

**Key words:** Distortion, Data reduction, Perceptual Model, Digital audio systems

**Ocjena nelinearnih digitalnih audio sustava pomoću smanjenja transparentnosti uporabom složenog ispitnog signala.** Ocjena nelinearnih digitalnih audio sustava pomoću smanjenja transparentnosti uporabom složenog ispitnog signala. Do sada poznate i primjenjivane metode mjerenja izobličenja pokazale su se neprimjenjive za otkrivanje grešaka u audio sustavima sa smanjenom brzinom prijenosa. Za ispitivanje sustava sa smanjenom brzinom prijenosa potrebno je koristiti signale koji su po svojim karakteristikama što sličniji stvarnim signalima, tj. govoru i glazbi. Istovremeno, pri ocjeni kvalitete potrebno je uzeti u obzir psihoakustičke procese koji se odvijaju u ljudskom uhu. Za potrebe ispitivanja sustava s redukcijom podataka razvijen je složeni ispitni signal čija kumulativna funkcija razdiobe amplituda slijedi Gaussovu razdiobu. Ulazni i izlazni signal transformiraju se iz objektivne domene razina u domenu subjektivne pobude. Mjera izobličenja je smanjenje transparentnosti sustava i definirano je kao omjer pobude signala greške prema pobudi ulaznog signala.

**Ključne riječi:** izobličenje, redukcija toka podataka, percepcijski model, digitalni audio sustavi

## 1 INTRODUCTION

An important division of digital audio systems is made taking into account their linearity, so they can be divided into linear and nonlinear, where the latter are often referred to as systems with reduced bit rates. Such a division is based on the relation between the bit depth of the system  $n$ , the sampling frequency  $f_s$ , and the bit rate  $BR$  of the digital output signal. For linear digital systems the bit rate is equal to the product of the sampling frequency and bit depth:

$$BR = f_s \cdot n \quad (1)$$

while for nonlinear digital systems the bit rate is lower than the value obtained by making such a product:

$$BR < f_s \cdot n \quad (2)$$

The introduction of digital audio devices into everyday use has evoked the need to develop appropriate measurements

to be applied during production, operation and maintenance of such devices. The attempts to use the measurements defined for analog devices [1, 2] have revealed certain faults due to the fact that specifics of digital audio systems have not been considered. The measurement procedures could only cover digital devices as complete units, while the measurements could not be made separately for the digital and the analog part of these devices. Therefore, specific measurements for digital systems have been defined, as well as the test signals and procedures to be used in such measurements [3-5]. Although any kind of signal can be implemented in these measurement procedures, a reference signal is required for comparison purposes. For these reasons, the proposed method of evaluation uses a signal whose statistical characteristics are similar to the ones obtained for speech and music, while maintaining full control of signal parameters. In addition, the reference signal is not needed for evaluation, since it can

be reconstructed from the known initial values of signal parameters.

## 2 MEASUREMENTS ON DIGITAL AUDIO SYSTEMS

The specific characteristics of digital devices, the types of errors and the origins of distortions demand that new measurement procedures have to be defined. Furthermore, the test signals have to be chosen in such a way that the measurements yield unambiguous results and reveal all possible anomalies of the measured device.

There are three important facts that determine the difference in behaviour between analog and digital systems during the measurements. The first one is the dependency of all frequency components in a digital system on the sampling frequency of the system. As a direct consequence, non-harmonic components can occur in the output signal as a result of intermodulation between the input signal and the components of the sampling frequency. The second fact focuses on avoiding the distortion caused by aliasing, i.e. overlapping of the basic spectrum with components from the lower sideband of the sampled signal. Consequently, all digital systems have a bandwidth limitation, which means that only test signals with frequency components lying below half the sampling frequency can be used. The third fact is related to quantization of the signal amplitude and the quantization noise as the by-product of this process, being the difference between the real value of the signal and the assigned numerical value after quantization. The spectrum of the quantization noise depends on the relative frequency relations between the input signal and the sampling frequency, unlike in the analog systems.

The basic types of errors that occur in digital audio systems can be divided in two groups. The first group contains the errors of input/output devices, while the second group consists of errors that occur during signal processing.

### 2.1 Errors of the Input/Output Devices

Errors of the I/O devices are present in all digital systems, regardless of the type of interfaces used for their interconnection (analog or digital). The use of A/D and D/A circuitry increases the probability that errors will occur, as opposed to the case when digital inputs and outputs are used. Nevertheless, working in the digital domain brings specific problems of its own.

Although the main goal of this paper is to determine the quality of nonlinear digital audio systems in which the signal processing errors are dominant, it is necessary to understand the origin of errors related to input and output devices in order to get a complete overview of a digital audio system. The analysis of measurements of I/O errors resulted in important ideas related to the choice of

appropriate test signals. Based on results of this analysis, the demands were defined that have to be fulfilled during the measurements of digital audio systems in order to gain credible results and unambiguous quality markers for the system under test.

All relevant characteristics of a measured digital audio system can be divided into six groups: the input characteristics, the output characteristics, system response, distortion, S/N ratio and inter-channel interference.

### 2.2 Signal Processing Errors

The development of digital audio systems with reduced bit rate (RBR systems) stressed the need for introducing new procedures for determining the quality in many fields of use, from transmission, recording and broadcasting in the professional part to receiving and recording in the part intended for home devices. Various directions of development and goals of final use, as well as a wide area of different usage require well defined ways of determining the quality in order to make comparisons and optimal choices for each use [6 - 8]. The expression "determination of quality" is used deliberately because distortions in the classical sense cannot be associated with RBR systems due to specific signal processing and algorithms used in such systems. Instead of measuring the distortions, an assessment is made of the overall auditory "impression" of a certain RBR system. In the early stage of development, the RBR systems were evaluated based on groups of specially trained listeners whose task was to determine the noticeability of signal change using a set of auditory tests. This was a time-consuming and expensive procedure not adequate for everyday control and checks of correctness. Moreover, such a procedure depended on a number of different factors (signal type, listening space, test method, level of training, listeners' musical preferences, etc.).

The basic structure of the ISO/MPEG encoder [9] is shown in Fig. 1 as the most widely used RBR system. It consists of:

1. Frequency-time mapping (filter bank) used for breaking down the input audio information into undersampled spectral components. Depending on the filter bank used in this stage, they are called sub-bands or spectral lines.
2. The output of the filter bank or a parallel transformation output used for calculating the instantaneous masking threshold according to the psychoacoustic model. This information determines the next function steps of the quantizing device and the encoder.
3. Quantization and encoding of sub-band samples or spectrum lines with the goal to keep the noise introduced by the quantization below the masking threshold level.

4. Forming of blocks that contain information about the quantized signal, information about the data sequence in the block, amplification factor of the block and additional information.

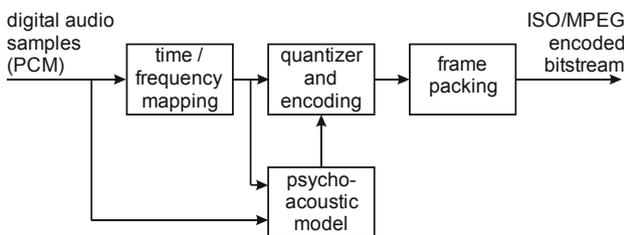


Fig. 1. Basic structure of an ISO/MPEG encoder

### 3 THE PROPOSED MEASUREMENT METHOD

Conventional measurement methods developed for analog systems mostly use stationary sinusoidal test signals and their combinations. A common characteristic of these signals is that the probability density distributions of their amplitude deviate significantly from the Gaussian distribution which is common for speech and musical signals. Moreover, the spectra of these test signals are time invariant although they can contain various frequency components. For this reason, it was often very difficult to ascertain a meaningful relationship between the values of distortion measured using conventional methods and subjective evaluation grades obtained from listening tests. Furthermore, in real life, audio systems are never used for transfer or recording of signals such as sinusoidal or rectangular. Digital audio systems with reduced bit rate are based on the principles of irrelevance, redundancy and entropy of the signal. It is not possible to give an overall evaluation of quality for a nonlinear digital audio system under test by measuring only the common parameters and distortion. Therefore, the use of transparency is proposed for evaluation of RBR audio systems. This measure evaluates the similarity of the output signal to the input signal of an RBR system, taking into account human hearing.

The initial idea in defining the new test signal was to use a simple waveform as the basis for building the final form of the test signal. An easy way of defining and modifying the test signal was achieved in this manner, while maintaining complete control over all parameters of the signal.

The examination of simplicity of generating and defining several types of waveforms has shown that triangle and saw signals are the optimal basic waveforms based on which the test signal can be defined. The spectrum analysis performed on these two types of waveforms reveals that the spectrum of triangular signal contains only odd harmonics, while the spectrum of the saw signal contains both

odd and even harmonics. Furthermore, saw signals have a more favourable energy distribution within their spectra. Specifically, the amplitude levels of the harmonics of triangular signals fall off at a higher rate than the amplitude levels of the saw signal harmonics. Taking into account these two findings, the saw signal was chosen as the basic building element for the new test signal.

As bandwidth limitation is necessary for all digital systems, the rate of change for the back edge of the saw signal has to be limited as well, thereby ensuring that the excitation signal is maintained within the frequency range of the system. For this reason, the saw signal is modified so that it never has a perfectly vertical back edge, but always a sloped one.

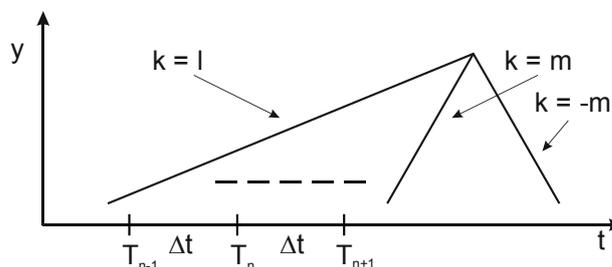


Fig. 2. Parameters of the new test signal

Figure 2 shows the basic parameters used for determining the waveform of the new test signal. The expression used for generating the signal can be defined as:

$$y_{T_n} = y_{T_{n-1}} + k \cdot \Delta t \tag{3}$$

and represents the defining equation for a straight line. The directional coefficient  $k$  determines the slope of the straight line, i.e. the rate of change of signal amplitude. As the signal must exhibit dynamically changeable characteristics, the directional coefficient is changed after every  $p$  periods of the signal that have been generated. In order to define the test signal as similar as possible to real signals in which the transient and the stationary parts of the signal have the same statistical probability of occurring, the directional coefficient  $k$  does not change in a linear fashion. Instead, the change of the directional coefficient is determined by a pseudorandom sequence of length  $m$ . The chosen parameter  $m$  defines a subset of natural numbers  $\{1, \dots, m\}$ . The pseudorandom sequence is generated in a way that each number from this subset will be used once, ensuring that there is no repetition inside the sequence. The back edge is defined with a straight line with directional coefficient equal to the highest value from the pseudorandom sequence, but with a negative sign. Although the length of the sequence  $m$  does not have a crucial influence on generation of the sequence when using a computer, sequences

of length  $m = 63, 127$  and  $255$  were taken into account due to simplicity of hardware implementation.

The increment of the signal is not determined from the values of the slope of the signal, i.e. the directional coefficients, but as the product of the increment coefficient and the quotient  $D/m$  where  $D$  is the dynamics of the saw signal.

The slope of the back edge determines the maximum excitation frequency in the baseband. This slope is determined from the number of samples the back edge contains, from which the ratio between the highest fundamental frequency and the sampling frequency is determined.

The slope of the back edge also defines the location of the first minimum in the signal spectrum. Thus, the basic range of excitation is defined for such a signal. The level difference between the spectral components contained in the basic excitation range and the higher order components is in the order of 40 dB, which corresponds to the energy distribution in the spectrum of real signals. Each slope of the front edge of the saw signal is repeated three times, which gives a frequency dependant duration of particular components.

The signal generated as described meets the following initial conditions: it is quite simple for generation, its spectrum can be easily modified and the signal parameters are chosen so that they enable controllable changes of signal characteristics.

Although the generated signal is highly variable over time, its amplitude distribution is rectangular and does not correspond to distributions of real signals. Therefore, the final implementation of the signal generator is shown in Fig. 3 and uses three synchronised generators, each of them working on the principle described above. The common characteristic for all three generators is that the pseudorandom sequences are mutually connected. The seed sequence is the same for all generators. The first generator yields a whole number coefficient  $k_{1i}$ . This coefficient is passed on to the first saw signal generator as the slope of the front edge. At the same time, this value is also used as a pointer to the seed sequence in the second generator. The obtained coefficient  $k_{2i}$  is then used in the same way for the second generator of the saw signal and as a pointer to the seed sequence in the third generator. An example of this procedure is shown in table 1. The first coefficient in the first pseudorandom sequence generator (PSG)  $k_{11}$  is equal to the first number in the seed, "3". At the same time it points to the 3rd position in the seed, number "2", making it the first coefficient  $k_{21}$  in the second PSG. Again, this coefficient also points to the 2nd position in the seed, number "1", making it the first coefficient of the third PSG. The same procedure is repeated in each subsequent time interval.

Table 1. Example of the CTS modelling for the seed sequence [3,1,2]

seed: [3,1,2]	time interval (i)			
	1	2	3	4
k1	3	1	2	3
k2	2	3	1	2
k3	1	2	3	1

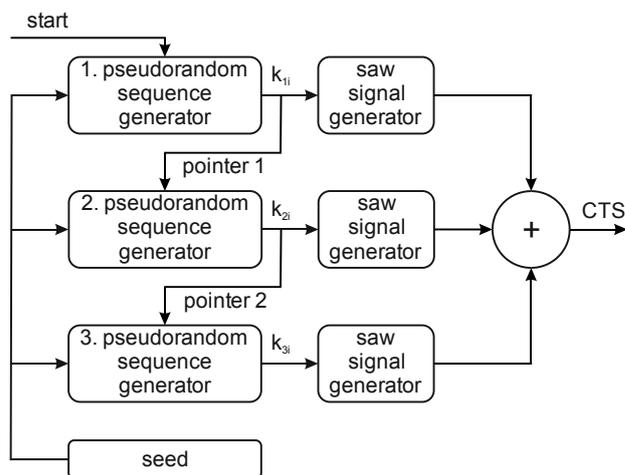


Fig. 3. Generation of the Combined Test Signal (CTS)

The amplitude distribution problem is solved by adding the signals obtained from all three generators. As shown in Fig. 4, the amplitude distribution of the test signal now follows the Gaussian distribution very well [10, 11].

Using a signal whose statistical characteristics are very similar to real audio information solves the first part of the problem. The evaluation of the amount of error must take into account the noticeability of this error by the end user by considering the processes that take place in the human

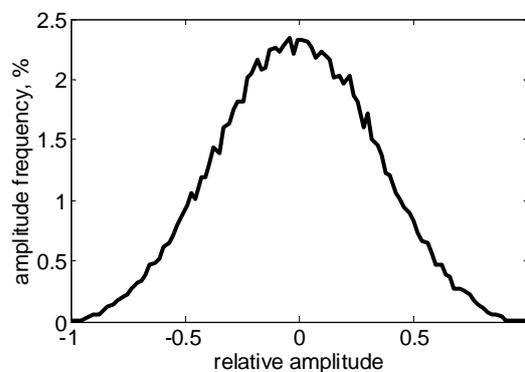


Fig. 4. Amplitude distribution of the Combined Test Signal in 100 bins

hearing system. The psychoacoustic effects shown and defined up to date are based on tests carried out with stationary stimulus, mostly sinusoidal signals or narrowband noise. The goal of these measurements is to use a procedure which will be based on already standardized parameters as much as possible, and that some space is left for upgrading based on new findings in the field of psychoacoustics, mostly on those connected with the dynamic behaviour of the human hearing system.

The proposed test setup which utilizes the combined test signal is shown in Fig. 5. The first phase in the evaluation of quality is the transformation of the original test signal from the objective domain of level values into the subjective domain of listening experience, i.e. into the stimulus domain. This transformation is done in two stages. The first stage involves the transformation from the linear frequency domain into the domain of the critical bands of the human ear, i.e. the Bark domain. In this stage signal levels are determined in each critical band. The real characteristics of critical bands are substituted with equivalent rectangular filters [12]. The addition procedure is described in ISO and IEC standards [13]. After the transformation in the frequency domain, the values of sound pressure obtained for all critical bands are transformed into levels expressed in decibels, (4).

$$L_{p_r} = 10 \log \frac{p_r^2}{p_0^2} = 10 \log \left[ \sum_i \left( \frac{p_i}{p_0} \right)^2 \right], \quad (4)$$

where the reference sound pressure  $p_0$  is  $20 \mu\text{Pa}$ ,  $p_r$  is the sound pressure in the corresponding critical band and  $p_i$  is the sound pressure of the  $i$ -th frequency line in that critical band. Then,  $L_{p_r}$  is the sound pressure level of the corresponding critical band.

The stimulus levels of the input signal  $L_{p_r}$  are then corrected by values which characterize the hearing threshold inside each critical band. The new corrected level inside the  $j$ -th critical band and  $i$ -th time interval is now marked as  $A_{ij}$ .

The sum of all values in the matrix which represents the input signal that has undergone the described transformation and correction is the total stimulus  $E_p$  for the original input test signal, (5). This transformation is made in the measuring instrument itself.

$$E_p = \sum_i \sum_j A_{ij} \quad (5)$$

The output signal from the system under test is undergoing the same procedure as the input signal, i.e. the transformation from the frequency domain to Bark domain and the correction regarding the average hearing threshold for each critical band. The total stimulus is not calculated for

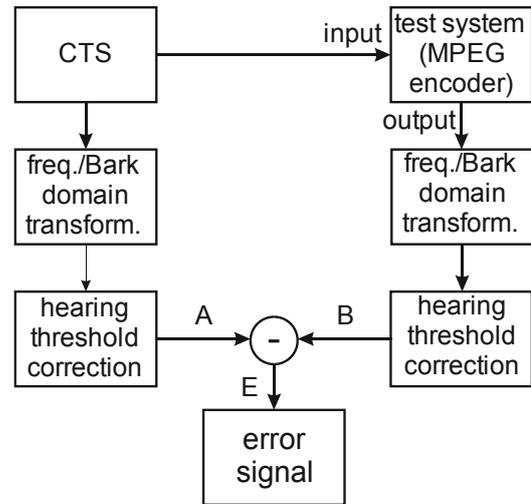


Fig. 5. Block diagram of the test setup which utilizes the Combined Test Signal

the output signal. Instead, the transform matrix  $\mathbf{B}$  of the output signal is subtracted from the transform matrix  $\mathbf{A}$  obtained for the original input signal. Thereby, the error signal matrix  $\mathbf{E}$  is obtained for the output signal in relation to the input signal as:

$$\mathbf{E} = \mathbf{A} - \mathbf{B} \quad (6)$$

The sum of absolute values of all elements contained in the error matrix  $\mathbf{E}$  is calculated as the total error stimulus  $E_g$  according to:

$$E_g = \sum \sum |E_{ij}| \quad (7)$$

where  $E_{ij}$  is the error in the  $i$ -th time interval and in the  $j$ -th critical band. The ratio of the error signal  $E_g$  and the total stimulus of the input signal  $E_p$  represents the transparency reduction  $K$  for the investigated system, (8).

$$K = \frac{E_g}{E_p} \text{ or } K\% = \frac{E_g}{E_p} \cdot 100[\%] \quad (8)$$

The values of transparency reduction defined in this manner lie between 0 and 1. The value 0 corresponds to lossless transfer of information. In the extreme case when there is no output signal, the error stimulus  $E_g$  is equal to the stimulus of the input signal  $E_p$ . In this case transparency reduction reaches the value of 1 or 100 %.

#### 4 MEASUREMENTS PERFORMED USING THE COMBINED TEST SIGNAL

All actual methods of quality evaluation defined for digital audio systems with data reduction implement the calculation of the ideal transformation for the known reference signal on the input and then compare it with the output

from the system for which the quality has to be determined. Such a procedure is practical and simple enough for application in laboratory conditions, but it is not possible to use it for field measurements such as quality evaluation of a radio broadcasting system.

The method of signal generation shown in Fig. 3 makes it very easy to meet the requirements stated for field measurements. In order to evaluate the transparency reduction, it is enough to know the seed sequence and to determine the delay between the generated signal and the signal obtained on the output of the system. The seed sequence can be prewritten in the read-only memory (ROM) of the measurement system. It can also be transmitted in the header of the digital audio signal, e.g. in the Broadcast Wave Format.

The exact time point in the output signal which matches the beginning of the input test signal must be determined by cross-correlating the generated and the received encoded signal, as shown in Fig. 6. The time point of the peak value of the cross-correlation subtracted from the length of the encoded signal represents the actual time delay introduced by the encoder.

$$R_{xy}(m) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n)y(n+m), \quad (9)$$

$$-\infty < m < \infty .$$

In (9),  $x(n)$  and  $y(n)$  denote two discrete signals. In this case these signals are the input and the output signal of the measured system. The results obtained with correlation of two signals defined with (9) are more unambiguous if the spectrum of signals is broader. The combined test signal defined as described earlier has a broad spectrum which is favourable for applying such a correlation.

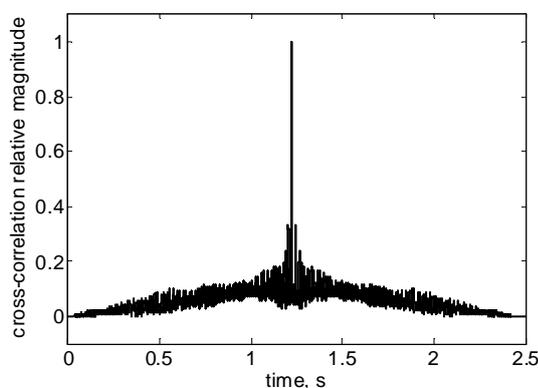


Fig. 6. Cross-correlation function between the original and encoded CTS signal (MPEG1 Layer III, 128 kbit/s)

Besides the cross-correlation between the input and the output signal, the autocorrelation of both the input and the output signal was also calculated as a tool for detecting the periodicity in each signal. When the combined test signal is used in measurements, the well defined and narrow maxima in the autocorrelation function provide the means for synchronization on the output of the system if there is no possibility to make a direct comparison with the original signal.

In order to verify the validity of the proposed measuring method and the combined test signal, software versions of the ISO/IEC MPEG 1 Layer I, II and III encoders and decoders were used, originally intended for checking and demonstrating the possibilities of making reductions in audio signals. The measurements were made using different sets of values chosen for relevant parameters of the combined test signal. Only single channel systems were investigated. The test results are shown for MPEG 1 Layer III as the most widely spread codec for audio today. Codec designates a system consisting of an encoder and a decoder. Transparency reduction was calculated and compared for all three layers of the MPEG 1 standard.

Figure 7 shows 3 ms long excerpts of the waveforms of the original CTS and the one encoded with MPEG 1 Layer III. The differences of the two waveforms are clearly visible, suggesting that a certain amount of error is introduced during the encoding process. Further evaluation reveals how much this difference will be noticeable for the end user, i.e. the listener. The waveform of the signal encoded with 128 kbit/s is not shown because the difference to the original signal cannot be observed, suggesting that the quality of the system at this bit rate is high.

Figure 8 shows the power spectra of the original CTS and the CTS again encoded with MPEG 1 Layer III encoder at two bitrates, 48 and 128 kbit/s. While the power spectrum of the signal encoded at 128 kbit/s does not differ

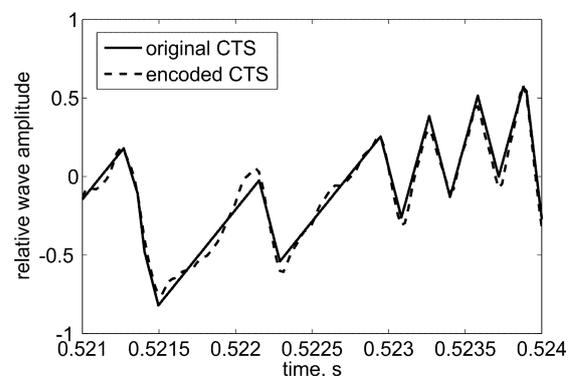


Fig. 7. The waveforms of the original CTS and the one encoded with MPEG 1 Layer III at 48 kbit/s

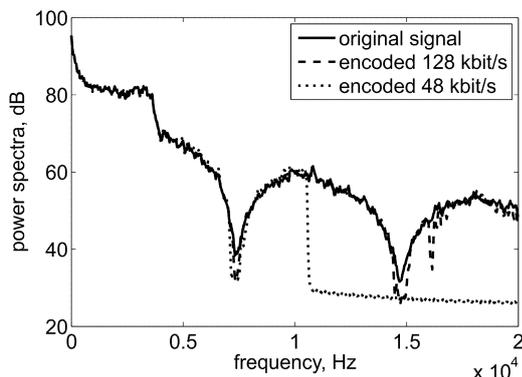


Fig. 8. Power spectra of the original CTS and the ones encoded with MPEG 1 Layer III at 48 and 128 kbit/s

significantly from the one calculated for the original signal, the power spectrum of the signal encoded at 48 kbit/s is visibly different in the high frequency region above 10 kHz.

The same phenomenon can still be observed after the transformation from the frequency domain to the critical band domain and the applied correction that takes into account the hearing threshold of the human hearing system. Figure 9 shows the power spectra of the same signals transformed to the critical band domain and corrected as described earlier.

Finally, Figs. 10 and 11 show the error signal matrix E as the difference between the transform matrices A and B obtained for the original and the encoded CTS signal. The CTS signal was again encoded with MPEG 1 Layer III at 48 and 128 kbit/s. In accordance with the already shown results, it is again visible that the amount of error is neg-

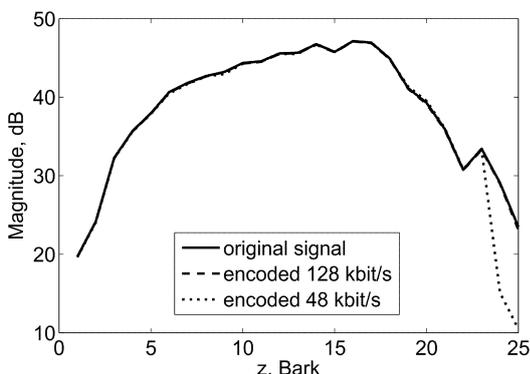


Fig. 9. Power spectra of the original CTS and the ones encoded with MPEG 1 Layer III at 48 and 128 kbit/s after the transformation to the critical band domain and correction taking into account the human hearing threshold

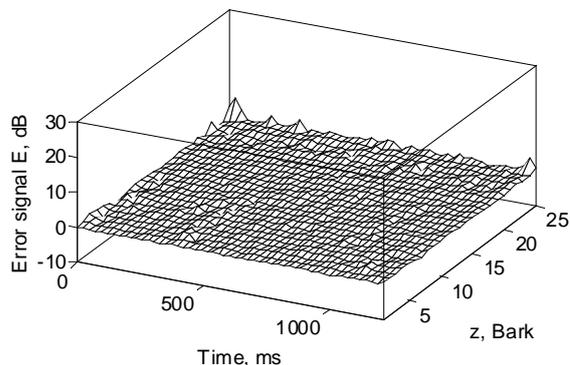


Fig. 10. Error signal E for CTS encoded with MPEG 1 Layer III at 128 kbit/s

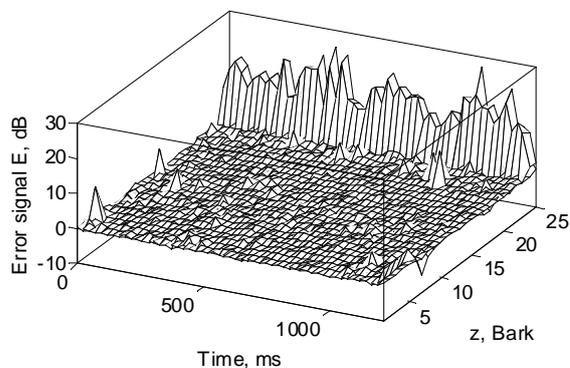


Fig. 11. Error signal E for CTS encoded with MPEG 1 Layer III at 48 kbit/s

ligible for the signal encoded with 128 kbit/s. This finding suggests that this particular encoder offers very good quality of encoded audio signal which is indistinguishable from the original signal. On the other hand, a bit rate of 48 kbit/s is inadequate for this particular encoder due to the high amount of error that occurs in the highest critical bands, as shown in Fig.10.

In order to compare all three codecs defined in the MPEG 1 standard, a transparency reduction analysis was made at different bit rates. Specifically, Layer I encoder was examined in the range from 32 to 448 kbit/s, Layer II from 32 to 128 kbit/s and Layer III from 48 to 128 kbit/s. The resulting transparency reduction indexes are shown in Fig. 12.

The results clearly reveal that the transparency of a system does not change linearly with the increase of bit rate, but that the improvement of quality becomes less and less significant after a certain bit rate threshold is reached. These results are in good correlation with the results of subjective listening tests made for chosen bit rates of the

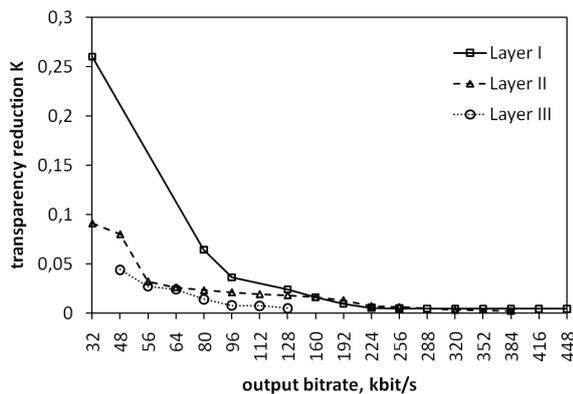


Fig. 12. Transparency reduction indexes obtained for CTS signal encoded with MPEG Layer I, Layer II and Layer III at various bitrates

ISO/IEC data stream, as well as the results of the tests made by other authors [14]. Furthermore, for a given bit rate the smallest transparency reduction was observed for MPEG 1 Layer III, especially at low bit rates, suggesting that this codec offers the best quality of encoded signal at a given bit rate or that the bit rate can be lower if a certain level of quality has to be obtained. This finding is consistent with the topology of the codecs themselves, with each new layer solving certain issues that were present in the past ones, thereby improving the quality of encoded signal.

## 5 CONCLUSION

Digital audio systems with reduced bit rate are based on the principles of irrelevance, redundancy and entropy of the signal. The amplitude distribution in time of real signals follows the Gaussian distribution and the spectra of speech and music signals are not stationary. The common measurement methods mostly use stationary sinusoidal signals or their combination in order to maintain the simplicity. The common characteristic of these signals is that the amplitude distribution deviates from the Gaussian distribution and the spectra of such test signals are constant over time, although they can contain different components.

It is not possible to make an evaluation of the subjective impression of quality of a digital system with reduced bit rate by measuring only the standard parameters and the usual types of distortion. For this reason, transparency is used to evaluate the quality of such audio systems. Based on the results of subjective listening tests, the information is gained on the similarity of the output signal to the input signal. In order to avoid lengthy and costly subjective listening tests, a new measurement signal was developed and a new method of quality evaluation was proposed.

The combined test signal consists of three saw signals all having the same slope of their back edges, while the slopes of their front edges are determined with the pseudorandom sequence. The length of this sequence can be changed and it determines the duration and the precision of the measurement. For short-time testing of transmission paths, the sequence with length equal to 63 is quite enough. For more detailed laboratory testing the sequence length can be extended up to 255 numbers. Each of the three saw signals is generated according to its sequence. All three sequences are mutually linked, thereby ensuring the repeatability of the test signal. This feature is especially convenient when there is no signal from the generator at the measurement location, e.g. during measurements of transmission or radio broadcasting systems.

The transparency reduction represents the proposed measure of quality of the system under test. It is obtained by implementing a nonlinear transformation of the test signal from the domain of objective level values to the stimulus domain, i.e. the domain of the subjective listening experience. This signal is used as a reference in the process of determination of transparency reduction for a given system. The measurements defined in this manner can be used on different types of encoders and decoders. The results of the proposed measuring procedure in this paper show good correlation with the results of subjective listening tests done with real listeners, as in [14]. In these tests, the quality of the low-bitrate codec systems increased with the rise of the data bitrate.

The CTS can be used also for evaluating low-bitrate multichannel audio systems. In this case, each channel has to be evaluated with a CTS generated with different seeds because some multichannel coders use joint stereo approach to reduce the bitrate. If the system was evaluated with the same CTS in each channel, the coder would recognize the test signals as identical and would encode only one channel, but with a higher effective bitrate which does not correspond to a real situation.

## REFERENCES

- [1] K. Jambrošić, B. Ivančević, I. Zorić, "Electroacoustic Measurements on Audio Coders", *Automatika*, Vol. 41, No. 1-2, pp. 63-67, 2000.
- [2] K. Crnković, M. Maletić, H. Domitrović, "Digital Audio Measurements for Broadcasting" *Proceedings of 17th International Conference on Applied Electromagnetics and Communications ICECOM*, pp. 139-142, 2003.
- [3] E. Benjamin, "Evaluating Digital Audio Artifacts with PEAQ", *113th AES Convention*, (San Francisco, USA), pp.1-10, 2002.
- [4] A. Kurittu, J. Samela, A. Lakaniemi, V. Mattila, N. Zacharov, "Application and Verification of the Objective

Quality Assessment Method According to ITU Recommendation Series ITU-T P.862”, *Journal of Audio Engineering Society*, Vol. 54, No. 12, pp. 1189-1202, 2006.

- [5] EBU-TECH 3309, “Evaluation of Cascaded Audio Codecs”, EBU Report, EBU, 2005
- [6] M. Šalovarda, I. Bolkovac, H. Domitrović, “Estimating Perceptual Audio System Quality Using PEAQ Algorithm”, *Conference Proceedings of 18th ICECOM*, 2005.
- [7] AES Staff, “Measuring and Predicting Perceived Audio Quality”, *Journal of Audio Engineering Society*, Vol. 53, No. 5, pp. 443-448, 2005.
- [8] A. Magadum, “Evaluation of PEAQ for the Quality Measurement of Perceptual Audio Encoders”, *Proceedings book of the 29th AES International Conference: Audio and Mobile and Handheld Devices*, (Seoul, Korea), pp. 5-2, September 2006.
- [9] K. Brandenburg, G. Stoll, “ISO-MPEG-1 Audio: A generic Standard for Coding of High-Quality Digital Audio”, *AES International Conference: Collected Papers on Digital Audio Bit-Rate Reduction*, (New York, USA), pp. 30-42, 1996.
- [10] I. Đurek, H. Domitrović, M. Maletić, “The Sensitivity Of Audio Distortion Measurement Methods”, *Proceedings of The First Congress Of Alps Adria Acoustics Association*, (Portorož, Slovenia), pp. 511-518, 2003.
- [11] H. Domitrović, I. Zorić, B. Ivančević, “The New Combined Test Signal for Evaluating Perceptual Codecs”, *Proceedings of the IMEKO TC-4*, (Naples, Italy), pp. 175 – 178, 1998.
- [12] E. Zwicker, H. Fastl, *Psychoacoustics: Facts and Models*, Springer, 1999.
- [13] ISO Standard 532:1975, “Acoustics – Method for calculation loudness level”, ISO, 1975
- [14] EBU-TECH 3296, “EBU subjective listening tests on low-bitrate audio codecs”, EBU Report, EBU, 2003



**Hrvoje Domitrović** received BSc, MSc and Ph.D. degrees from University of Zagreb, Faculty of Electrical Engineering and Computing in the field of electroacoustics. His doctoral thesis in 1997 was on the objective grading of perceptual audio systems. From 1982 to 1986 he had been employed at the Nikola Tesla Company as an R&D engineer in the field of digital speech transmission. From 1986 to 1991 he worked at Croatian Radio as a chief engineer responsible for implementation of digital audio technology in

broadcasting. Since 1991 he has been with FER, where he is currently vice-dean. He teaches acoustics and audio. His research areas are architectural acoustics, noise control and digital audio broadcasting. In 1993 he is one of the four owners of Croatian AES Section, acting as a chair from 1997 to 2000. He is member of AES, ASA, EAA. He is collaborating member of the Croatian Academy of Engineering.



**Kristian Jambrošić** received his B.Sc., M.Sc. and Ph.D. degree (in the field of psychoacoustics) from the Faculty of Electrical Engineering and Computing, University of Zagreb in 1996, 2000 and 2004, respectively. Since 1997 he is employed at the Department of Electroacoustics of the same faculty, first as a research assistant on

two research projects, and currently as an Assistant Professor. He teaches a number of undergraduate, graduate and doctorate courses in the field of acoustics, audio systems and electronics. His scientific interests are aimed towards building acoustics, noise protection and psychoacoustics. From 2008 he is the senior researcher on the national research project “Acoustic characteristics of composite elements in civil engineering”. He is also a MC member of the European COST project TD0804, board member of the Acoustical Society of Croatia, and general secretary of the European Acoustics Association.



**Marko Horvat** received his B.Sc. and Ph.D. degrees from the University of Zagreb, Faculty of Electrical Engineering and Computing in Zagreb, Croatia in 2001 and 2010, respectively. He is currently employed as a research assistant at the same faculty and is actively involved in both research and teaching in the fields of noise, product sound quality, room and building acoustics, sound reinforcement and electroacoustics. He has published more than 30 scientific papers and was involved in the proceedings of numerous scientific congresses. He is a member of Acoustical Society of Croatia and of Audio Engineering Society.

scientific congresses. He is a member of Acoustical Society of Croatia and of Audio Engineering Society.

#### AUTHORS' ADDRESSES

**Prof. Hrvoje Domitrović, Ph.D.**

**Asst. Prof. Kristian Jambrošić, Ph.D.**

**Marko Horvat, Ph.D.**

**Department of Electroacoustics,**

**Faculty of Electrical Engineering and Computing,**

**University of Zagreb,**

**Unska 3, HR-10000, Zagreb, Croatia**

**email: hrvoje.domitrovic@fer.hr, kristian.jambrosic@fer.hr**

**marko.horvat@fer.hr**

Received: 2010-10-31

Accepted: 2010-11-11