UDC 004.825 **doi: 10.20998/2522-9052.2018.4.03**

O. Nesmiian

Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine

METHOD OF THE MULTITONAL SIGNAL CONSTRUCTION IN DECISION SUPPORT SYSTEMS ACS

Subject: The article deals with the issues of creation of a psychoacoustic model of perception of sound quality of sound signals, the development on its basis of the method of creating a multi-tonal signal, which provides a high correlation with subjective evaluation. **Purpose:** to develop a multitoon signal creation method. **Objective:** To analyze the distortion of the sound signals that arise during reproducing them on the sound reproduction devices. Classify these distortions and assess the extent of their impact on subjective assessment. Allocate the most significant of them in terms of impact on the quality of sound. Investigate modern methods for assessing the quality of reproduction of sound signals, standards and recommendations devoted to this task; to evaluate the effectiveness of modeling the human auditory system used in the development of these methods; formulate the scientific and applied requirements for developing objective quality assessment method. **Conclusions:** The application of the developed method will allow obtain results that coincide with the data of subjective and statistical examinations (SSE) with the accuracy required for practice.

Keywords: DSS; ACS; multitonal signal; a model of quality perception.

Introduction

Formulation of the problem: Despite the fact that today there are a fairly large variety of processing audio signals algorithms, as well as elements of sound reproduction systems and their characteristics, there is a problem of choosing the best combination of sound reproduction devices and algorithms for the audio signal arise. Since the audio reproduction devices cause large distortions in the reproduced signal itself, existing methods for evaluating their quality do not provide a comprehensive characterization of the distortion, but only, at best, indicate a high degree of difference of the reproduced signal from the original. All this makes the actual task of developing fundamentally new methods for creating multi-tone signals, which allow to ensure the efficiency and reliability of the solution of functional problems by operators.

Analysis of recent research and publications: The problem solving of quantitative assessment of sound quality began to be engaged in the middle of the last century. In the 40s-50s, the work of H. Fletcher [1], D.K. Gannet [2] and WBSnow [3], which investigated the required bandwidth for quality reproduction of certain musical instruments, male and female voices; sensitivity of the human auditory system to changes in the reproduced frequency range, perception of nonlinear distortions, etc. The disadvantages of these works include the lack of completeness of the perception of distortions in the case of their simultaneous presence in the evaluated signal and the impact on the subjective quality assessment. In the 1990s, several methods for objectively assessing sound quality using psychoacoustic modeling were developed. In the magazine AES 23 (Audio Engineering Society) were published works, which considered the problem of objective quality assessment using knowledge of the human auditory system features. As a result of the received data, the method [4] PERCEVAL (PERCEptual EVALuation) was proposed. This method based on the comparison of the signal of additive noise with the threshold of masking, calculated from the reference signal and the calculation of the probability of detecting these noise by the human auditory system. The next generation of this method was the method PAQM (Perceptual Audio Quality Measure) [5], which is based on the comparison of "internal representations" calculated for the reference signal and the investigated signal.

The result of comparing the "internal representations" of signals is interpreted in the assessment of the sensitivity of distortions. In addition to the aforementioned methods, a number of reports were also presented at the AES conferences on psychoacoustic methods for assessing the quality of sound [6], [7], [8]. As a result of a detailed investigation and comparison of all available methods, the PSQM algorithm was approved in recommendation P.861. The results of objective assessments obtained with the PSQM method have a high correlation (> 97%) with the results of subjective listening tests, which allowed the successful implementation of this method in the telecommunications market. Since the original algorithm proposed by the ITU-T can not be used to assess the quality of sound paths in general (taking into account losses, delays and distortions), the authors continued to work on improving the method in order to realize the possibility of taking into account time delays. As a result, extensions of the standard were developed: PSQM +, PSQM99 and PSQM / IP, which it was decided to integrate into the new PESQ (Perceptual Evaluation of Speech Quality) standard. This standard was presented as a new recommendation of the ITU-T Rec. P.862 [9]. The second group of ITU-R developed a method that allows an objective assessment of the quality of broadband music codecs. As a result of the research, the group presented the recommendation ITU-R BS.1387 [9] describing PEAQ (Perceptual Evaluation of Audio Quality). Currently, this method is the most reliable tool for evaluating the quality of musical signals.

Purpose of the article: The purpose of the article is to develop a method for creating a multi-tone signal in the DSS ACS.

Main material presenting

The magnitude and spectrum of nonlinear distortions introduced by the sound tracks of the reproduction devices depends on the reproduced signal characteristics.

Consequently, the generated multi-tone signal must have a similar spectral composition with the original signal, i.e. envelope of the spectrum of the artificial signal should correspond to the envelope spectrum of the input signal at each instant of time.

The proposed method for obtaining a multi-tone signal in its base is based on the audio signal processing procedures used in the psychoacoustic model used in the MREO-1 180 / 1EC 11172-3 [10] standard.

Outbound audio signals used to subjectively assess the sound quality of reproducing devices are subdivided into samples, the size of which depends on the amplitude-time characteristics of the signal. The size (duration) of the samples is determined by the resolution of the human auditory system and specified by the requirements of implementing the FFT algorithm.

It is established that the auditory system of a person processes audio signals of a duration from 80 to 140 ms [11]. For implementation of FFT algorithms are used samples, the number of readings in which are equals two in the integer degree. So, at a standard sampling frequency of 44.1 kHz, the largest sampling length not exceeding 80 ms and corresponding two raised to the integer power, is a sample of 2048 signal counts, that is, approximately equals 48 msec. In some cases, the length of the sample may vary depending on the structure of the audio signal within the sample, at moments of sharp jumps, the sound sample length will be 1024 counts (24 msec).

Change the length of the sample criteria is a peak factor value [12]. Psychoacoustic entropy [13] is the criteria for compression algorithms for digital audio data, which is calculated for a long sampling. The use of shorter samples (12 msec) is inappropriate, since the human hearing apparatus, due to inertia, is not able to accurately assess the spectral composition of short sound segments, i.e. to notice the additive noise.

The use of a 50% overlay of samples used in the short-term spectral analysis allows to reduce the distortions that arise during orthogonal transformation.

Before the FFT procedure, the sound samples should be processed by the Hannah function, which allows to clear the signal from additional distortions, known as the Gibbs effect, that arise when performing a direct FFT.

As source material for investigation were created multi-tone test signal which are used in the ICAO radio exchange

The main steps of the algorithm for creating a multi-tone signal:

Step 1. Analysis of the sound signal in the time domain for the selection of samples with energy above the threshold.

Step 2. Calculate the peak factor of the signal of the selected samples, which is necessary to select their length (1024 or 2048 counts of the sound signal).

Step 3. Calculation of the normalizing factor.

Step 4. The calculation of the energy spectrum using FFT is performed for each selected sample.

Step 5. Calculate the energy of the sample signal in critical audible bands; selection from the spectrum of the sampling signal of the tonal components.

Step 6. The list of noise-like components creation.

Step 7. Thinning of the spectrum of tonal and non-tonal (noise-like) components considering the absolute threshold of sensitivity and simultaneous masking.

Step 8. Adjustment of the energy of the spectral components to ensure the equality of energy of the initial and synthesized multi-tone signals.

Step 9. Restore the time form of the multi-tone signal.

On the first step of creating a multi-tone signal, an analysis of the output signal in the time domain is performed.

During creating a multi-tone signal it makes no sense to analyze all the samples in a row, but rather select those whose energy level exceeds a certain threshold.

On the second step for decision making about the length of the sample, the peak factor value calculation for a long sample is performed using the formula (1):

$$CF(dBFS) = \\ = \frac{s_{\max}(dBFS)}{\sqrt{\sum_{i=1}^{N}\left(s(n)(dBFS) - s(dBFS)\right)^2 \big/ N}}, \quad (1)$$

where i - input signal sample number,

n - input signal count,

$s_{\max}$ - maximum reference value,

s - average reference value for the sampling signal.

All values are expressed here are in $dBFS$. At magnitude of the peak factor greater than $dBFS$, the length of the sample for analysis is 1024 (short sampling), otherwise it is equals 2048 counts (long sample) of the sound signal.

On the third step of creating a multi-tone signal, the rationing of the fac coefficient and rationing of the counts selection of the input signal is made:

$$fac = 10^{L_p/20} \big/ Norm \quad (2)$$

$$x(n) = fac \cdot s(n), \quad (3)$$

where $x(n)$ - normalized samples.

In formula (2), a tone signal with a frequency of 989 Hz and an amplitude of 0 dB is used to determine the normalization coefficient $Norm$, and the spectral maximum for several frames is calculated. The value of this maximum is taken as the value of the normalizing factors, and then the calculation of the energy spectrum of the sample is performed:

$$X(k) = 10\ln\left|\frac{1}{N}\sum_{n=0}^{N-1} h(n)\cdot x(n) e^{-jkl2\pi/N}\right|^2 + \quad (4)$$

$$+90.309dB,$$

where $k = 0,1...N/2$;

$N = 2048 \; or \; 1024$;

k - index of the spectral component of FFT;

$h(n) = 0.5 \cdot \left[1 - \cos\left(2\pi n / (N-1)\right)\right]$ – window function Hannah;

$X(k)$ – level of k spectral component, in dB.

In this case, the normalization of FFT coefficients to 96 dB SPL is used.

On the fourth step, the energy of the sample signal is calculated in critical listening lanes [14]. The total energy of all spectral components in each critical band is calculated by the expression (5):

$$E_{sb}(z) = 10\lg\left(\sum_{k \in z} 10^{X(k)/10}\right), dB \qquad (5)$$

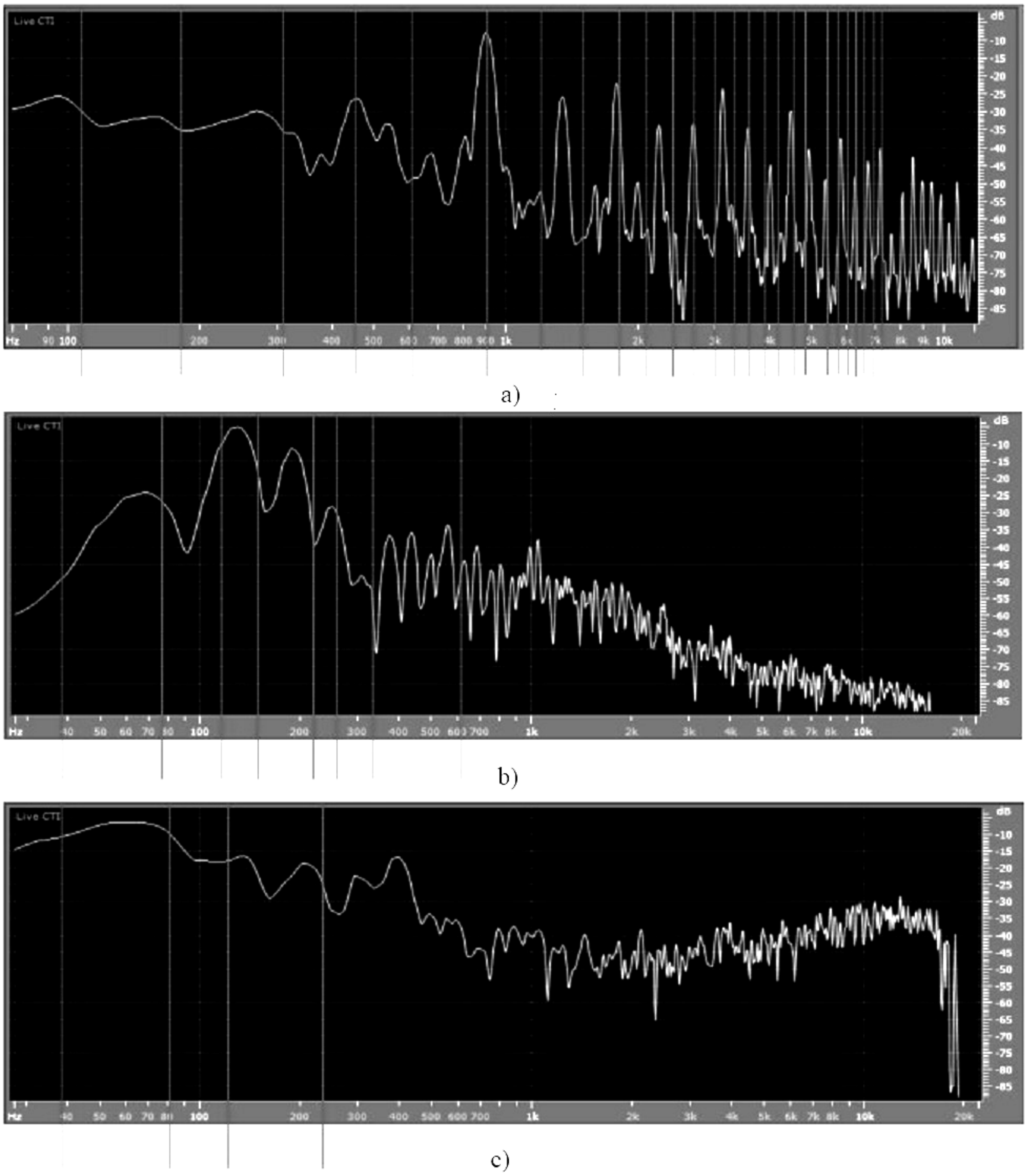
In the fifth step, the allocation of so-called tonal components is carried out. Spectral components of the energy spectrum $X(k)$ exceeding the frequency, except for two neighboring at least 7 dB, are identified as tonal. Note that the area of the examined frequencies expands with increasing frequency [10]. The set of tonal spectral components of 8T is determined by the following expression

$$S_T = \left\{ X(k) \middle| \begin{matrix} X(k) \triangleright X(k \pm 1), \\ X(k) \triangleright X(k \pm \Delta_k) + 7dB \end{matrix} \right\}, \qquad (6)$$

where $X(k)$ – k spectral maximum,

Δ_k depends on the bandwidth and varies from 2 to k.

Examples of spectra of samples of a sound signal with different number of tonal components are shown in Fig. 1.



a)

b)

c)

Fig. 1. Spectra of samples of signals:
a – 26 tonal, b – 8 tonal, c – 4 tonal

At the sixth stage, the formation of a list of noisy components is taking place.

To do this, from the output signal of the spectrum of the sample, all sub-tone and adjacent components are removed, after which the energy $X_{nm}(z)$ of the spectral components in each critical band is calculated.

As a result, the real noise-like components of the sample signal are replaced by one spectral component whose energy is equal to the total energy of the noise-like components that have fallen into the critical listening band (7):

$$X_{nm}(\bar{k}) = 10\lg\left[\sum_{j=l}^{u} 10^{0.1 \cdot X(j)}\right], \quad (7)$$
$$\forall X(j) \notin \{X_{lm}(k, k \pm 1, k \pm \Delta k)\}.$$

At the seventh stage, the thinning of the received tonal and noise-like components is carried out. That is mean, that the number of components is reduced based on two criteria: the absolute threshold of sensitivity and frequency degradation. The tone and noisy components below the absolute threshold of sensitivity are discarded, and only the components that satisfy the condition are left

$$X_{lm,nm}(k) \geq X_y(l) \quad (8)$$

where $X_{lm,nm}(k)$ - the sound pressure level corresponding to the absolute threshold of sensitivity in silence for the k spectral component.

Then, using a sliding window with a width of 0.5 barks, from any pair of dedicated spectral components appearing in this window, one that has the highest level of energy is selected.

At the eighth stage, there is an adjustment of the energy remaining after the reduction of the spectral components in order to ensure the equality of the energies of the initial and generated signals in each critical listening band.

To do this, in each critical listening band, the scaling factor is calculated as the ratio of the output signal energy to the sum of the energies of the generated multi-tonal test signal components:

$$scf(z) = \frac{E_{sb}}{\sum_{k=l}^{u} X_{mt}(k)} \quad (9)$$

where $X_{mt}(k)$ – components of the multi-tone signal in the critical band.

After that, the adjusted energy of the component is calculated:

$$\bar{X}_{mt}(k) = X_{mt}(k) \cdot scf(z), \forall k \in z. \quad (10)$$

This correction is necessary to achieve the same level of load on the audio path of the tested device, real and synthesized multi-tone signals.

At the final stage, the recovery of the time form of the multi-tonal signal is carried out using the inverse Fourier transform.

At the end of the multi-tone synthesis procedure, a passport that contains information on sample length, sets of numbers (indexes) of the spectral components, and information about the type of each component (tonal or noisy) is created. For the convenience of analysis of the recorded signal (to search the beginning and end of the content of the signal), before the main multi-tone signal and after it, is added special tone pulses.

For analyzing of the listening space noise between the tone pulses and the useful signal, a section of the signal with a zero level is inserted.

Conclusion

On the basis of a thorough analysis of psychoacoustic models, procedures are selected and validated for assessing the magnitude of certain types of distortion introduced into the signal by sound tracks of sound reproducing devices.

The revision of these procedures due to their joint use within the framework of a single psychoacoustic model was performed.

The method of creating an artificial multi-tone signal for precise allocation of the nonlinear distortion (ND) signal and its processing after passage through the sound path of the tested device for analysis and evaluation of the electromagnetic radiation is developed.

In the developed method, the calculation of a set of quantitative characteristics of distortions of different types, which have the greatest impact on the generalized assessment of sound quality.

The following list includes the following indicators:

the likelihood of hearing ailment of the ND;

the number of strongly distorted fragments of the signal;

noise mask ratio in three frequency bands;

effective dynamic range;

peak factor;

distinction of specific volumes of signals in frequency bands;

uneven frequency response in the field of voice frequencies.

For formalization of the transition from a multidimensional quality indicator to a generalized objective quality assessment by one number, a multiple regression model is used.

References

1. Fletcher, H. (1942), "Hearing the determing factor for high fidelity transmission", *Proceedings of the IRE*, No. 30(6), pp. 266-277, DOI: https://doi.org/10.1109/JRPROC.1942.230995.

2. Gannet, D.K. and Kerny, J. (1944), "The discernibility of changes in program band width", *Bell Systems Technical Journal*, Vol. 23, pp. 1-10.

3. Stuart, J.R.: (1992), "Implementation and measurement with respect to human auditory capabilities", *AES UK DSP Conference*, London, September 1992.

4. Paillard B. Mabilleau, P. and Morissette S. (1992), "PERCEVAL: Perceptual Evaluation of the Quality of Audio Signals", *Journal of the Audio Engineering Society*, 1992 January/February, Vol. 40 (1/2), pp. 21-31.

5. Beerends, J.G. abd Stemerdink, J.A. (1992), "A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation", *J. Audio Eng. Soc*, Vol 40, No. 12. pp. 963-973.

6. Geddes, E.R. and Lee, L.W. (2003), "Auditory Perception of Nonlinear Distortion – Theory", *AES 115th Convention*, Paper 5890.

7. Olive, S. (2004), "A Multiple Regression Model For Predicting Loudspeaker Preference Using Objective Measurements: Part I-Listening Test Results", *116th Convention*, 2004 May 8-11, Berlin, Germany.

8. Olive, S. (2004), "A Multiple Regression Model for Predicting Loudspeaker Preference Using Objective Measurements: Part II - Development of the Model", *117th Convention*, 2004 Oct 28-31, San Francisco, USA.

9. ITU-T Recommendation P.862 (2001), PESQ an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, February 2001.

10. ISO/IEC 11172-3 (1993), Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s - part 3: Audio, 1993.

11. ITU-R Recommendation BS.562-3, Subjective assessment of sound quality, 1990.

12. Shitov, L.V. and Belkin, B.G. (1970), "Statistical characteristics of signals representing natural sound, and their use in the research of electro-acoustic systems", *Works of the Research and Development Film Photo Institute*, No. 56, pp. 77-173.

13. Kovalgin, I.A. and Vologodin, E.I. (2004), "Tsifrovoe kodirovanie zvukovyih signalov " [Digital coding of sound signals], Korona print, Saint Petersburg.

14. Sheluhin, O.I. and Lukiantsev, N.F. (2000), *Digital processing and voice transmission*, Radio and communication, Moscow, 456 p.

Надійшла (received) 30.09.2018
Прийнята до друку (accepted for publication) 07.11.2018

Метод створення мультитонального сигналу
в системах підтримки та прийняття рішень АСУ

О. Ю .Несміян

Предмет: У статті розглянуті питання створення психоакустичної моделі сприйняття якості звучання звукових сигналів, розробки на її основі методу створення мультитонального сигналу, що забезпечує високу кореляцію з суб'єктивною оцінкою. **Мета:** розробка методу створення мультитонального сигналу. **Завдання:** Провести аналіз спотворень звукових сигналів, що виникають при відтворенні їх звуковідтворюючими пристроями. Класифікувати ці спотворення і оцінити ступінь їх впливу на суб'єктивну оцінку. Виділити найбільш значущі з них з точки зору впливу на якість звучання. Провести вивчення сучасних методів оцінки якості відтворення звукових сигналів, стандартів і рекомендацій, присвячених даній задачі; оцінити ефективність моделювання слухової системи людини, що застосовується при розробці цих методів; сформулювати наукові та прикладні вимоги, що пред'являються до методу об'єктивної оцінки якості, який розробляється. **Висновки.** Застосування розробленого методу дозволить отримати результати, що збігаються з даними суб'єктивно-статистичних експертиз (ССЕ) з необхідною для практики точністю.

Ключові слова: СППР; АСУ; мультитональний сигнал; модель сприйняття якості.

Метод создания мультитонального сигнала
в системах поддержки и принятия решений АСУ

О. Ю. Несмиян

Предмет: В статье рассмотрены вопросы создания психоакустической модели восприятия качества звучания звуковых сигналов, разработки на ее основе метода создания мультитонального сигнала обеспечивает высокую корреляцию с субъективной оценкой. **Цель:** разработка метода создания мультитонального сигнала. **Задача:** Провести анализ искажений звуковых сигналов, возникающих при воспроизведении их звуковоспроизводящими устройствами. Классифицировать эти искажения и оценить степень их влияния на субъективную оценку. Выделить наиболее значимые из них с точки зрения влияния на качество звучания. Провести изучение современных методов оценки качества воспроизведения звуковых сигналов, стандартов и рекомендаций, посвященных данной задачи; оценить эффективность моделирования слуховой системы человека, применяется при разработке этих методов; сформулировать научные и прикладные требования, предъявляемые к методу объективной оценки качества, который разрабатывается. **Выводы:** Применение разработанного метода позволит получить результаты, совпадающие с данными субъективно-статистических экспертиз с необходимой для практики точностью.

Ключевые слова: СППР; АСУ; мультитональний сигнал; модель восприятия качества.