Increasing the Accuracy of Signal Formation by Changing the Sampling Rate

Oleg Popov
MTUCI
Moscow, Russia
olegp45@yandex.ru

Andrey Borisov
MTUCI
Moscow, Russia
a.borisov199@gmail.com

Tatiana Chernysheva
MTUCI
Moscow, Russia
krba2012@yandex.ru

Kirill Orlov
MTUCI
Moscow, Russia
kvororloff@yandex.ru

Valentin Abramov
MTUCI
Moscow, Russia
vabramov@mail.ru

Abstract—The problem of changes in the audio signal when changing the sampling frequency arose a long time ago, but interest in it has recently become more acute with an increase in the quality of signal formation and transmission. To some extent, this was also determined by the methods of assessing the quality of changes in the harmonic signal without taking into account changes in the actual broadcast signal. Algorithms for the formation of a new digital sequence after changing the sampling rate provided perfect quality in the conversion of the sinusoidal signal and large errors in the conversion of the real signal. Algorithms for changing the sampling rate Fd have been improved.

The aim of this work is to study the features of changes in the sampling frequency in the paths of the broadcast channel and to develop algorithms for converting the sampling frequency in the frequency domain with an increase in conversion accuracy.

It is shown that the algorithms for converting the sampling rate in the time domain have reached perfection and cannot be improved, so the transition to processing in the frequency domain is important. An algorithm for converting the sampling frequency in the frequency domain has been developed. The spectrum of the error signal for various types of converters is determined and the need for its compensation is determined. A method for introducing predistortions for converters with known and unknown sequence of sample rates used has been developed.

The use of the results of this work makes it possible to improve the quality of broadcasting signals and information programs.

I. Introduction

The problem of reducing the distortion of the sampled of sound broadcasting signal (SBS) arose simultaneously with the creation of digital recording and broadcasting systems. At the first stage, the change in sampling frequency was carried out through an analog signal. With each conversion, quantization noise and hard-to-remove distortions of amplitude and phase-frequency characteristics were added. Evaluation of the quality of conversion, in accordance with existing methods, is carried out using a stationary harmonic signal, which It is converted with little or no distortion. On the real signal and the conversion predictions are quite large. In the future, algorithms and devices for digital conversion of the sampling frequency were created, which introduced significant distortions into the signal. Algorithms have been improved, and currently, with modifications in converters, they do not exceed a couple of hundred quantization steps, and for some algorithms, just a dozen.

Thus, algorithms for sample rate conversion in the time domain have reached a level of perfection and are unlikely to be improved further. However, sample rate conversion is widely used in analog-to-digital conversion, particularly with increased sampling rates. The inaccuracies in generating discrete values for the new digital sequence reduce the overall quality of the digital system. The goal of this study is to investigate the characteristics of sample rate conversion in broadcast channel systems and develop algorithms for frequency domain sample rate conversion with improved conversion accuracy. Solving this problem can be achieved through research and evaluation of distortions introduced into the audio signal by existing methods of sample rate conversion. Significant reduction in distortions is possible by transitioning from time-based to frequency-based sample rate conversion methods. Spectral parameters of the broadcast signal were used to accomplish this task. Computer modeling was also employed as a basis for conducting research on sample rate conversion in the time and frequency domains.

The proposed material discusses the possibilities of the original conversion algorithm, which makes it possible to reduce the distortion of the conversion to several quantization steps, and in the presence of information about the sequence of changes in the sampling frequency to one step.

II. Sample rate conversion algorithm.

The basic sampling rates used in the broadcast channel paths were:

- in the studio (program generation path-PGP) 44.1 kHz (CD) and 48 kHz (digital tape recorders, sound engineer's console);
- for signal distribution systems in the country (primary distribution path-PDP) 32 kHz (sound broadcasting), 31.25 kHz (TV sound);
- for systems for bringing the signal to the listener (secondary distribution path-SDP), for example, in digital broadcasting systems of 32 and 24 kHz [1].

Numerous encoders with a compact input and input signal representation still use pulse code modulation with standard sampling rates (F_d).

When changing F_d, the values of the discrete signal in new time positions are synthesized using interpolation, of varying complexity. An example of the implementation of such an algorithm for multiple sampling rates, 48 – 32 kHz, is shown in Fig. 1.

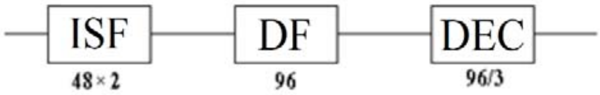


Fig. 1. Block diagram of the transformation of F_d

In the first block - an increase in the sampling frequency (ISF), upsampling, there is an increase in F_d due to the addition of zero samples, In the digital filter (DF) interpolator, zero samples are filled with interpolated values. In the decimator block (DEC), part of the samples is discarded.

A simplified field interpolator is shown in Fig. 2. With the help of delay lines, multipliers, and an adder, zero samples are filled with interpolated values. Even for a harmonic sine wave, regulatory compliance is achieved with around 1200 delay lines.

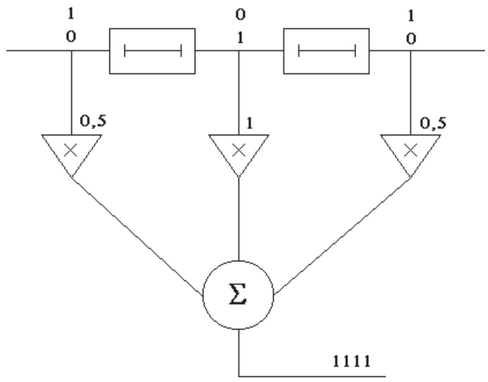
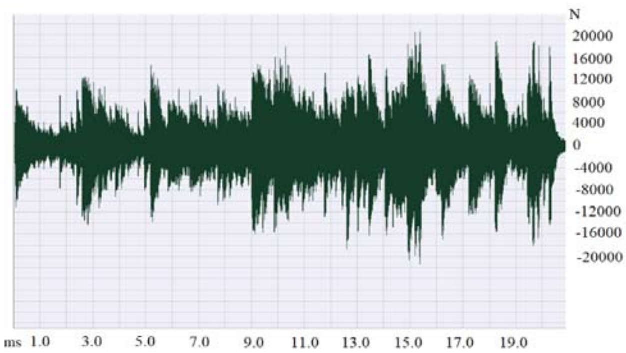


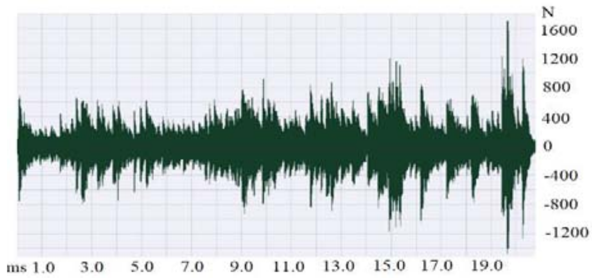
Fig. 2. First-order interpolator filter

In the absence of multiplicity between F_d, several steps of increase F_d are used until the error due to the time inaccuracy does not exceed the quantization step, or a given inaccuracy.

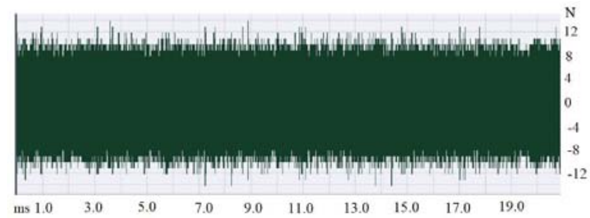
When conducting subjectively statistical tests (SST), distortion of the sound signal is especially noticeable against the background of the sound of the piano, which at the first stage was used as a test signal [2,3]. In Figures 3a, 3b, 3c, 3d shows the oscillations of the original signal and the error signal, which was obtained by subtracting from the original signal with a sampling rate of 44.1 kHz, the converted to a signal with a sampling rate of 48 kHz and vice versa.



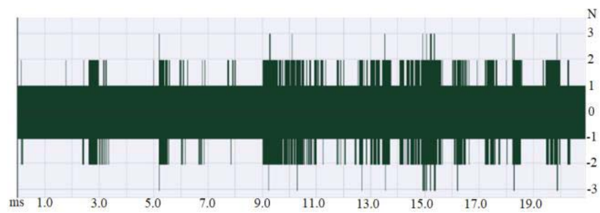
a) An excerpt of the sound of the piano



b) Error signal after changing F_d 44.1-48-44.1 kHz in Adobe Audition editor (old version)



c) Error signal after changing F_d 44.1-48-44.1 kHz in the Audacity 23 editor



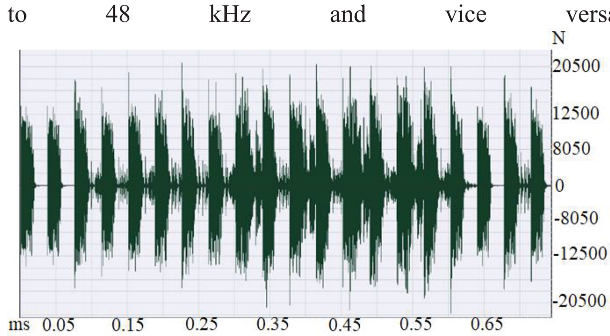
d) Error signal after changing F_d 44.1-48-44.1 kHz in on-lint audio converter 23

Fig. 3. Oscillograms of the original signal and the error signal after sampling frequency conversion

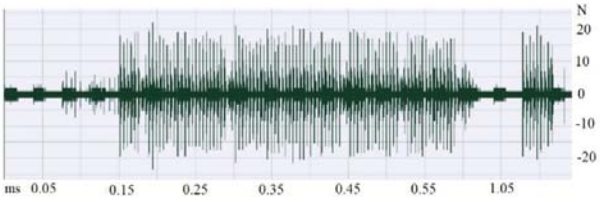
As we can see, the algorithms for converting F_d successfully cope with their task on the harmonious sound of the piano and having sections of long-term quasi-stationarity. The situation worsens when you try to convert F_d to the sound of jazz or any sound with a large number of noise components.

As we can see from the above waveforms, errors are especially large at the moments of changing sound objects, i.e. at moments of non-stationarity on the "attacks" of sound. Recall that such distortions are especially noticeable to the listener, because. A change in the nature of the signal rise is noticed by the listener already at a duration of 0.3 ms. [1]

Especially large errors arise when converting a sound signal containing both staccato and legato sounds. In Fig. 4a, 4b) shows the oscillograms of such a signal (drums, vocals) and the error signal generated after converting F_d from 44.1 to 48 kHz and vice versa.



a) Excerpt of the sound of "drums and vocals"



b) Error signal after changing F_d 44.1-48-44.1 kHz the Audacity 23 editor

Fig. 4. Oscillogram of the “drums vocal” signal and the error signal after conversion F_d 44.1 – 48 – 44.1 kHz

In most sampling rate conversion algorithms, interpolation of varying complexity in the time domain is used to generate discrete samples at a new sample rate [4]. A further increase in the accuracy of the sample rate conversion is possible when switching to the frequency domain. A method for converting the sampling frequency (CSF) is proposed, for which patent RF No. 2405262.

The patent suggests representing a signal as a set of coefficients of discrete Fourier transform (DFT) (Fig 5), then adding or discarding the necessary number of zero samples to modify Fd. Afterward, inverse discrete Fourier transform (IDFT) is performed to generate a digital sequence at the new sampling frequency. Samples are added or discarded in the range between the original sampling frequency and the upper frequency of the transformable signal. With the proper selection of a window function, no signal energy should be present here, ensuring a reduction in distortion of the Fd transform.

In Fig. 5 on top of the X-axis is the spectrum of the original signal, under the axis of the transformed.

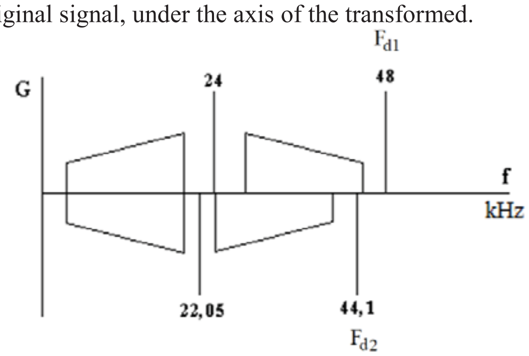
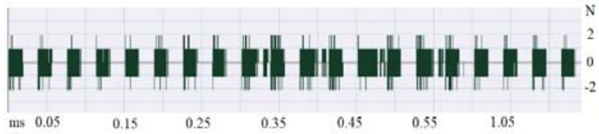
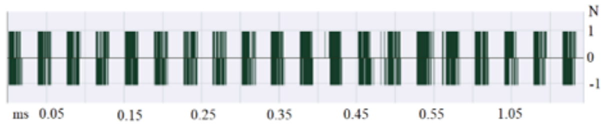


Fig. 5. Illustration of how to change the sample rate

To reduce the error of converting the sample rate in the case of a previously known sequence in the transmission or processing channel, the introduction of negative feedback (NF) allows. For example, a transmission channel uses a sequence of sample rates (Hz): 48,000 (studio), 32,000 (long-distance channel), 48,000 (radio house in another city), or 44,100, 48,000, 44,100 Hz in studio processing of files recorded on a disc. In this case, an error signal is formed at the transmitting end with such a sequence of transformations and is introduced into the transmitted signal in the form of predistortions [5]. The distortion that occurs in the digital signal at the output of this artificial circuit will be of the same nature as in the digital signal at the output of the real channel. After isolating these distortions, they are inverted and added as a pre-emphasis to the main digital signal at the input of the communication channel. Fig. 6a, 6b shows the oscillations of the error signal after converting 44,100 – 48,000 – 44,100 Hz using the developed algorithm and with the addition of predistortions. As we can see, the conversion errors can be reduced to one quantization step.



a) Oscillogram of the error signal of the converter of the CSF, the amplitude of the error up to 3 steps of quantization



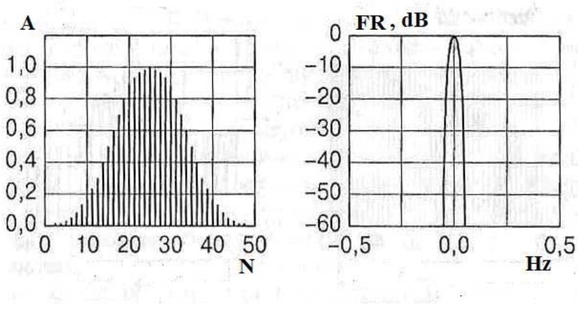
b) Oscillogram of the error signal of the converter of the converter with the introduction of predistortions, the amplitude of the error up to 1 step of quantization

Fig. 6. Oscillograms of error signals for the developed converter

III. FEATURES OF THE CONVERSION ALGORITHMS F_D.

When used in the conversion process of the signal at its presentation and in the frequency domain, as a rule, the accuracy of the conversion of the sample rate is determined by the accuracy of the FFT conversion of the long sample and the window function used.

The main requirement for the window function is the minimum spread of signal energy across the spectrum. We have quite successfully used the Nattol window in many similar developments, Fig. 7a, 7b.



a) 51-point windowof Nattol b) Frequency characteristics

Fig. 7. 51-point Nuttoll window and its frequency response.

The disadvantage of the window is that it is possible to maintain the unitandthe transmission coefficient at any percentage of overlap, therefore, after the IDFT, it is necessary to use an additional computerwitha window function. The main lobe for this function is w and re than for the commonly used ones, but in this problem this is not essential and can be compiled by a long sample.

For the one selected above, error value in quantitation levels (QL) depending on sample length (n) are shown in Fig. 8.

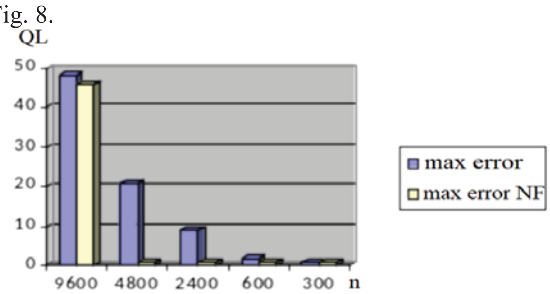
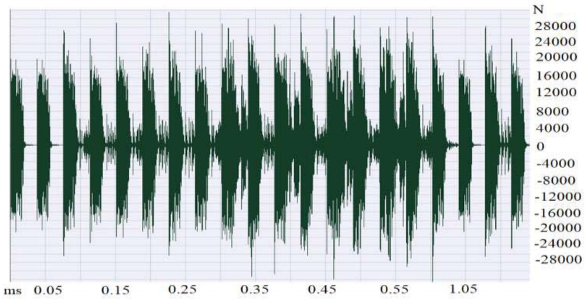
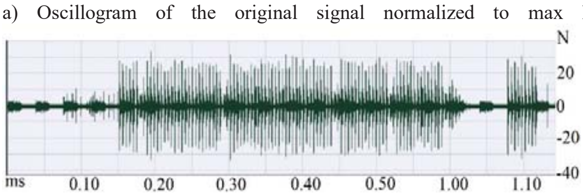


Fig. 8. Maximum amplitudes of the error signal (ER) generated with NF and without, depending on the number of discrete samples.

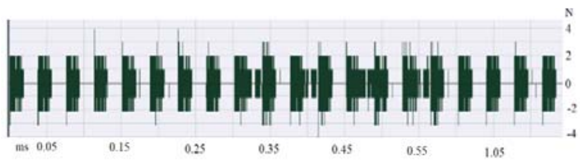
These error values were obtained for a signal whose maximum levels did not exceed -6 dB. The fact is that most sampling rate conversion algorithms are very sensitive to distortion limitations. When changing the sampling frequency, especially when increasing it, the peak value of the signal approaches the true value and usually increases [6,7]. To reduce such distortions, sound engineers, as a rule, underestimate the signal level. In Fig. 9a, 9b, 9c, 9d shows the oscillogram of the signal normalized to the maximum level and the oscillogram of the error signal.



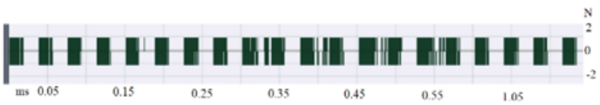
a) Oscillogram of the original signal normalized to max level



b) Oscillogram of the error signal for the maximum amplitude signal, amplitude up to 60 quantization steps.



c) Oscillogram of the error signal of the converter of the CSF, amplitude up to 3 steps of quantization.

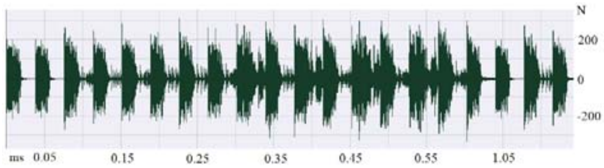


d) Oscillationlogram of the error signal of the converter of the CSF with the introduction of predistortions, the amplitude is up to 1 stepof quantization.

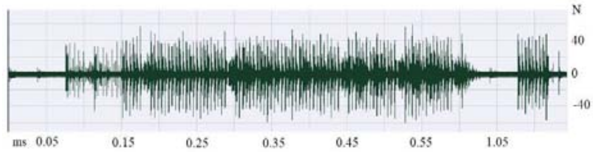
Fig.9. Oscillogram of a signal with a maximum level and an error signal, maximum values of about 30 quantization steps for an audio converter, 3 steps for the developed converter and 1 step for the same with negative feedback.

As we can see, the magnitude of the error for the "audio converter" converter, as well as for most others, depends on the amplitude of the signal and is not always linear [8,9]. Fig. 10a, 10b, 10c, 10d shows the initial signal and error signals: for the audio converter (error up to 60 quantization steps), for the developed algorithm and the algorithm with negative feedback.

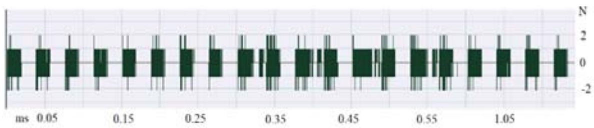
As we can see, distortions of oversampling depend on the amplitude of the signal and increase both in the case of an increase in the signal and in the case of its decrease. The developed algorithm is free from this drawback.



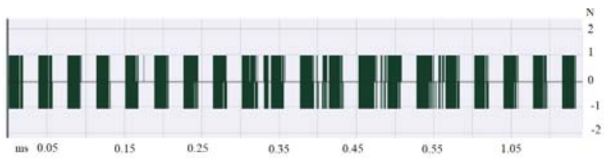
a) Oscillogram of the original signal, maximum amplitude of 320 quantization steps.



b) Oscillogram of the error signal of the online "audio converter", amplitude up to 60 quantization steps



c) Oscillatorogram of the error signal for the developed CSF algorithm, amplitude up to 3 quantization steps



d) Oscillationlogram of the error signal for the developed CSF algorithm using predistortions, amplitude up to 1 stepa quantization

Fig. 10. Signal with a level of -40 dB and error signals: for the audio converter (error up to 60 quantization steps), for the developed algorithm and the algorithm with negative feedback, the error has not changed

Fig. 11 shows histograms of errors normalized relative to the peak value for different signal levels.

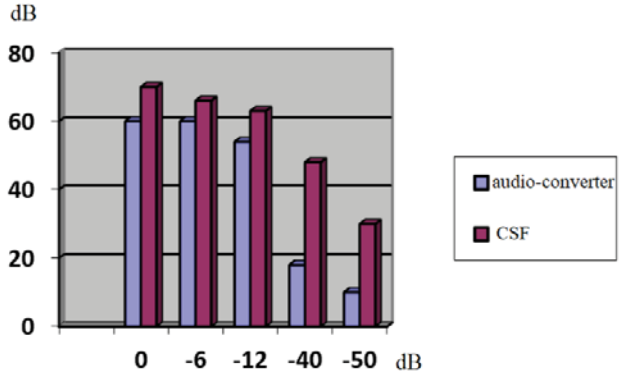
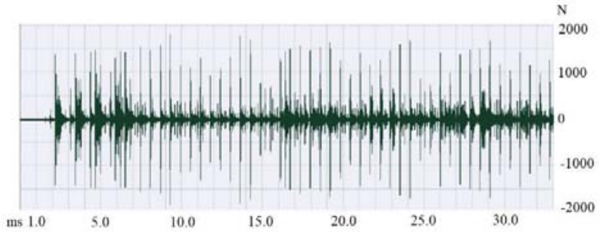


Fig. 11. Protection from conversion noise for the on-line converter "audio-converter" and converter CSF.

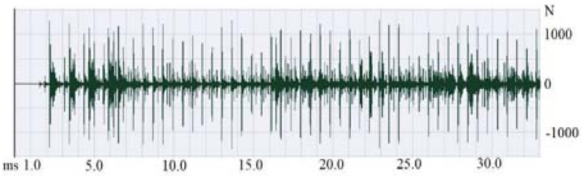
As can be seen from the results of the measurement of protection against oversampling errors for recording or transmitting a signal in accordance with the recommendations to provide at least 56-60 dB, this requirement can only be met for signals in the range from 0 to -40 dB. And taking into account the reduction of the signal level at the input by sound engineers by 6-10 dB and in an even narrower range [10,11].

IV. Changing the Sampling Rate and Signal Spectrum.

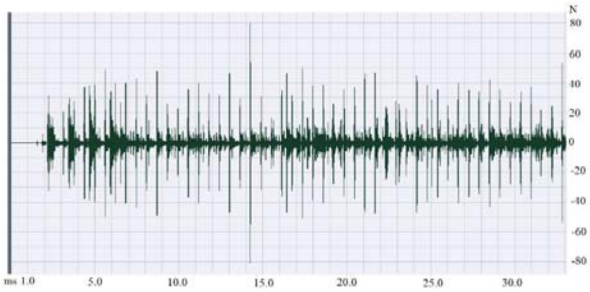
In practice, the sampling rate and the bandwidth of the transmitted signal are usually converted. For example, the long-distance path-the radio house - the secondary distribution to the transmitter. The sampling frequency varies from 32,000 to 48,000 Hz, while limiting the bandwidth of the transmitted frequencies from 20 to 15 kHz. In this case, filtering distortions are added to the distortions determined by the conversion process itself. As in the previous case, we will use the estimate of the error formed as the difference between the original and the signal of the passed path with sampling rates of 32-48-32 kHz. Fig. 12 shows the results of calculating the error signal for the Audacity 23 audio editor, the online converter "audio converter" and the converter CSF, including with feedback.



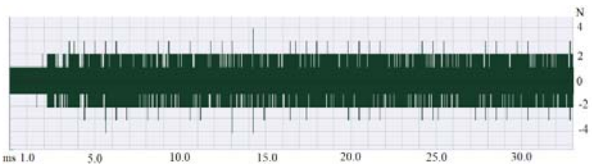
a) Error signal for Audacity audio editor, amplitude 1600



b) Error signal for online converter amplitude 1100



c) Error signal for the converter of CSF, amplitude 70



d) Error signal for CSF with NF, amplitude 4

Fig. 12. Error waveforms for four sample rate envelopes

As can be seen from the above measurement results, protection against distortion is clearly insufficient. In many cases, such modulation noises, so, arising in the background of the signal and in the signal spectrum, are neglected, as they are effectively masked in the auditory analyzer due to frequency masking. For example, when transmitting a signal in the MP3 format, the encoding noise is below the signal level by 30 dB. In a studio analog tape recorder, the protection against modulation noise is 35-40 dB, which does not reduce the quality of signal reproduction. Unfortunately, in this case, the spectrum of the error signal is not masked in any way by the spectrum of the signal itself, Fig. 13, 14. This is largely determined by the quality of filtering when changing the sample rate.

Error signal spectra for most converters. currently in use, are similar in nature to that shown in Fig. 12a, b, c, d, e, with a characteristic rise in the region of high frequencies.

The same Fig. 12 shows the spectrum of the error signal for the developed CSF algorithm. Judging by the spectrum, the error signal will be minimally noticeable against the background of the signal.

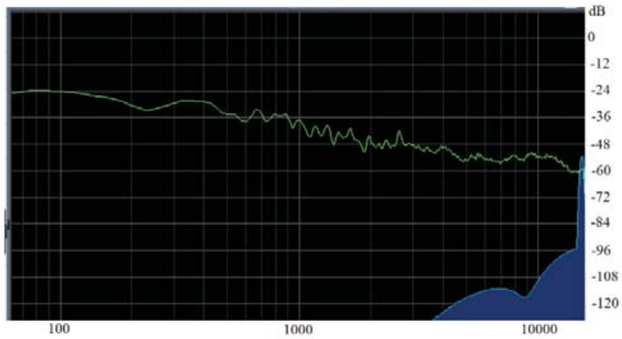


Fig. 13. Amplitude frequency response of the signal spectrum (upper curve) and error signal for the online converter (lower curve)

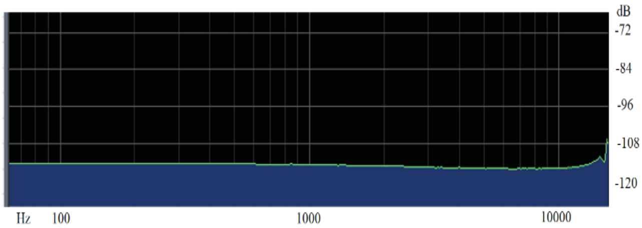


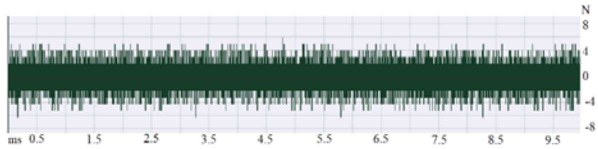
Fig. 14. Amplitude frequency response of the spectrum of the signal (upper curve) and error signal (lower curve), as well as the converter CSF

In some cases, it is not possible to generate a pre-distortion signal, since it is not supposed to return to the original sample rate signal.

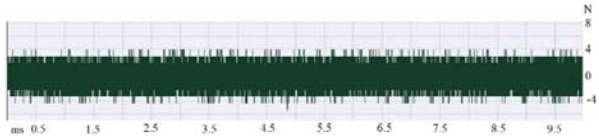
The effect of the pre-emphasis signal on the accuracy of signal formation with a new sampling rate was tested. A harmonic sinusoidal signal with a frequency of 440 Hz and an amplitude of -6 dB (test signal in Fig. 15a) at a sampling frequency of 44100 Hz was used as a test signal.

According to the CSF algorithm, this signal has been transformed into a signal at a sampling frequency of 48,000 Hz, and a test signal of 440 Hz, formed at a frequency of 48,000 Hz, has been subtracted from it. The error signal is shown in Fig. 15b. Then the original converted to a sampling rate of 48000 Hz was again converted into a signal from a sampling rate of 44100 Hz and a pre-emphasis signal was formed from it by subtracting from the original. In Fig. 15c shows the oscillogram of the error signal using pre-emphasis.

a) Oscillogram test signal



b) Oscillogram witherror ignal for converter according to the CSF algorithm error amplitude of 6 quantization steps



c) Oscillogram with ignala error for the converter according to the CSF algorithm with negative feedback, the amplitude of the error is 4 steps of quantization.

Fig. 15. Oscillograms of test signal and error signals when using the converter algorithm.

Based on the results of measurements on simple and complex test signals, it was confirmed that it is possible to increase the accuracy of the formation of a discrete signal at a new sampling frequency with a greater accuracy in reducing the amplitude of the error signal by 30%.

In addition to objective assessments of the quality of conversion using the developed CSF algorithm, subjective statistical tests were carried out on a variety of sound material, which confirmed the effectiveness of the algorithm [12].

V. CONCLUSION

To date, distortions, or rather changes in the signal in the transmission channel of the sound broadcasting signal are very large and are not covered by existing methods for measuring the objective parameters of the signal.

This also applies to algorithms for changing the sampling rate, the effectiveness of which is evaluated by the quality of the conversion of a stationary harmonic signal. Almost all algorithms successfully cope with this task, but problems arise when processing a real signal. You can understand sound engineers, in every possible way avoiding changes in the sampling rate on the sound material before processing it.

Changing the sampling frequency has become a widely used operation in every analog-to-digital conversion that involves oversampling. Inaccuracies in generating discrete values of the new digital sequence reduce the overall quality of the digital system.

Changing the sampling frequency has become a widely used operation in every analog-to-digital conversion that employs oversampling. The inaccuracy in generating the discrete values of the new digital sequence lowers the overall quality of the digital system.

Time-domain sampling rate conversion algorithms have reached perfection and can hardly be improved. The transition to frequency-domain processing is possible when ensuring the evaluation of coefficients in both sequences within the same frequency range band of the signal, which is defined as Fd/N, where N is the length of the sample. In processing, the integer value of the number of discrete samples and, consequently, the number of coefficients of the DFT should be maintained. For example, with a sample duration of 10 ms for sampling rates of 44.1, 48, and 32 kHz, integer samples would be 441, 480, and 320 samples, or the same number of coefficients of the DFT field.

The second limitation is related to the belief that FFT is only possible for sequences with a number of discrete samples that is a multiple of 2^n.

By now, a sufficient number of algorithms have been developed for FFT of sequences that are not multiples of 2^n.

The use of window functions with minimal side lobes also determines the feasibility of the algorithm.

Thus, as a result of the work done, there were:

- an algorithm for converting the sampling frequency in the frequency domain has been developed;
- an algorithm has been developed for converting the sampling frequency in the frequency domain with an increase in accuracy due to the introduction of predistortions into the signal;
- the accuracy of the formation of the converted signal for different sample lengths was determined;
- areas of undistorted conversion level for various types of converters is determined;
- the spectrum of the error signal for various types of converters is determined and the need for its compensation is determined
- A method for introducing predistortions for converters with unknown sequential sample rates has been developed.

In the acts of primary formation - studios, where it is necessary to maintain objectively high quality, it is simply necessary to assess the quality to preserve the shape of the real signal.

The study once again showed the insufficiency of the existing metrological support of the channels for transmitting and storing the sound broadcasting signal (SBS). Quality assessment is still carried out by changing the shape of a low-component, harmonic stationary signal, when, like almost all modern channels, this form is not preserved, providing high subjective quality for the listener.

Changes in the SBS in the transmission channel are large and are not covered by existing methods of objective rationing. Test signals, as a rule, stationary and low-component are not distorted in transmission systems with the elimination of redundancy, in contrast to a quasi-stationary multicomponent real signal.

The only method of quality assessment is subjectively statistical tests, which are not applicable if it is necessary to quickly assess the quality of the transmission and change it, for example, by means of new methods of encoding or audio processor processing. There is a need to investigate and identify the parameters of the signal, the changes of which determine the assessment of the quality of the transmission by the listener. It is necessary to involve sound engineers and musicians in such work, providing the ability to predict the assessment of sound quality by the listener based on the results of changes in the objective parameters of the SBS.

The use of the results of this work allows for improving the quality of broadcast signals and thereby increasing the popularity ratings of informational programs, which determine the economic efficiency of radio stations.

LITERATURE

[1] Y. A. Kovalgin, Digital Radio Broadcasting: Systems and Technologies. Hotline - Telecom. 2021, 580 p.

[2] M. Balik and O. Raso, "Assessment of Audio Signal Noise Reduction Based on Psychoacoustic Model," 2018 25th International Conference on Systems, Signals and Image Processing (IWSSIP), Maribor, Slovenia, 2018, pp. 1-5, doi: 10.1109/IWSSIP.2018.8439333.

[3] Hess, Richard L. (2018-10-17). "An overview of optimizing signal-to-noise ratio primarily in analog audio tape recording" (PDF). Aurora, Ontario, Canada. Archived (PDF) from the original on 2021-05-08. Retrieved 2021-05-07. (26 pages)

[4] T. Bai, L. Xie, Z. Li, J. Yang, Z. Chen and P. Wan, "A High-Precision Audio Z-Δ D/A Converter," 2020 IEEE 14th International Conference on Anti-counterfeiting, Security, and Identification (ASID), Xiamen, China, 2020, pp. 120-123, doi: 10.1109/ASID50160.2020.9271769.

[5] B. Ming and P. Wu, "Research on Audio Signal Denoising and Simulation Processing," 2019 International Conference on Communications, Information System and Computer Engineering (CISCE), Haikou, China, 2019, pp. 192-194, doi: 10.1109/CISCE.2019.00051.

[6] S. J. Schlecht, L. Fierro, V. Välimäki and J. Backman, "Audio Peak Reduction Using a Synced allpass Filter," ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, Singapore, 2022, pp. 1006-1010, doi: 10.1109/ICASSP43922.2022.9747877.

[7] P. Záviška, P. Rajmic, A. Ozerov and L. Rencker, "A Sur-vey and an Extensive Evaluation of Popular Audio Declipping Methods," in IEEE Journal of Selected Topics in Signal Processing, vol. 15, no. 1, pp. 5-24, Jan. 2021, doi: 10.1109/JSTSP.2020.3042071.

[8] V. A. Abramov, O. B. Popov, V. G. Orlov “Method and apparatus for companding audio broadcast signals”, RU Patent 2 691 122 C1, Jun. 11, 2019.

[9] Popov Oleg Borisovich, Abramov Valentin Alexandrovich, Prokurat Gleb Andreevich Compander system for transmitting the sound broadcasting signal of the Kompander 1 software. Version 2022. Certificate of state registration of a computer program No. 2022666583, Russia. Registration date: 05.09.2022.

[10] V. A. Abramov, O. B. Popov, V. G. Taktakishvili, A. A. Ovchinnikov “Method and device for measuring the power and steepness of the rise of areas of nonstationarity of acoustic signals”, RU Patent 2 731 339 C1, Sep. 01, 2020.

[11] V. A. Abramov, V. G. Taktakishvili, A. A. Ovchinnikov, O. B. Popov "Objective Assessment of Transmission and Informativeness of a Speech Signal According to Statistical Parameters" Proceedings of the FRUCT 24, Moscow, Russia, 8-12 April 2019, p. 741 – 746.

[12] Recommendation ITU-R BS. I 284-2 (01/2019). General methods for the subjective assessment of sound quality. BS Series. Broadcasting service (sound).