Analyses of Decimation Filter Stopband Attenuation Influence on Subjective Quality of Audio Signals

Nikola Petrović, Miloš Bjelić, Jelena Ćertić, Miodrag Stanojević, and Dragana Šumarac Pavlović

Abstract — **Decimation is a frequently used technique in audio signal processing. This paper investigates the influence of the values of stopband attenuation of decimation filters on the subjective perception of resulting audio signals. In practice, the demands for stopband attenuation of these filters are quite rigorous. The paper examines how justified these demands are from the psychoacoustic point of view. Basic procedure within such a research is a subjective test. Aside from objective parameters, the subjective evaluation is a method that can give insight into the characteristics of audio signal processing under testing. Results of subjective testing of the influence of different values of stopband attenuation on the perceptibility of interfering signal are presented for various audio signals. Two types of interfering signals were used in the subjective test. To increase the validity of psychoacoustic test, the probability of false detection of differences is measured by presenting subjects with equivalent signals. The results of subjective testing can prove to be useful in the design process of decimation filters for audio system applications.**

Keywords — **Subjective testing, Decimation, Filters, Psychoacoustics, Placebo, Critical bands.**

I. Introduction

SAMPLING frequencies in modern audio input devices are often set to a value that is a multiple of 48 kHz. For that reason, decimation is needed to convert the sampling frequency to 48 kHz. In recording and preprocessing of a

Nikola Petrović, School Of Electrical Engineering, University Of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia, (e-mail: pn130163d@student.etf.bg.ac.rs).
Miloš Bjelić, School Of Electrical Engineering, University Of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia, (e-mail: bjelic@etf.rs).
Jelena Ćertić (corresponding author), School Of Electrical Engineering, University Of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia, (e-mail: certic@etf.rs).
Miodrag Stanojević, School Of Electrical Engineering, University Of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia, (e-mail: miodragstanojevic@bitprojekt.co.rs).
Dragana Šumarac Pavlović, School Of Electrical Engineering, University Of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia, (e-mail: dsumarac@etf.rs).

signal with increased sampling frequency, nonlinearities of the system can introduce signal components in the non-audible part of the frequency spectrum. Decimation is a procedure realized as a cascade consisting of low-pass filtering and downsampling. Downsampling process results in aliasing of the input signal spectral components above half of the new sampling frequency, to the baseband of the output signal. To reduce this undesirable effect, input signal is filtered with a low-pass filter prior to downsampling [1]. In addition, the sigma-delta A/D convertors inherently use the oversampling-and-decimation signal processing chain. In sigma-delta A/D conversion decimation filter has the role of suppressing quantization noise. Noise shaping in sigma-delta A/D conversion translates quantization noise to a high frequency band, above audible frequencies [2], [3].

The measure of suppression of unwanted spectral components is not uniquely defined. Higher values of decimation filter stopband attenuation increase the complexity of the system. In this context, audio systems are an especially sensitive area, where rigorous demands need to be satisfied. In audio systems, particularly those that are implemented in the area of music production, undesirable effects must be minimized in the entire audio signal chain, ending with the listener. Therefore, when designing a decimation filter, a common request is for a relatively high stopband attenuation. The goal in this paper is to evaluate the validity and significance of these demands from the psychoacoustic point of view, i.e. subjective perception of a man as the end user of the system.

Large stopband attenuation criteria have large filter orders as a consequence, thereby increasing the computational complexity of decimation process. This paper does not focus on the details of the filter structure itself. FIR filter is used with the stopband attenuation as the only important design parameter.

The general design complexity of decimation filters with high stopband attenuation is the main motivation for investigating actual performances of these components in audio systems. In the fields of acoustics and audio technology measurements and objective physical parameters are often unable to provide a comprehensive view of systems performance, and the effects of certain system parameters on the output results [4]. In these situations, psychoacoustic tests are performed. The subjective perception is introduced as a qualitative evaluation of the system performance. Subjective testing was performed with a group of subjects where interference

perceptibility in resulting audio signal was tested. This paper presents the results of the improved test setup compared to our previous work on the evaluation of the influence of stopband attenuation of decimation filter on subjective perception of the audio signal [5]. The research was improved by increasing the number of subjects, using two types of high frequency interference signals, and introducing a more adequate signal preparation, thereby increasing the validity of conclusions drawn. Loudness levels of the original signals were equalized and the reproduction order was randomized. A set of control A-B pairs was introduced to detect the limit of the subjective detection of the differences in the processed signals. Furthermore, signals that proved inadequate in previous tests were removed. Signals used in this experiment were of various types, hence general conclusions can be drawn. Subjective testing was performed using two types of interference: a wideband signal and a narrowband signal. The results of these subjective tests can prove to be useful in decimation filter design stage.

The paper is organized in the following manner: the second section presents the methods for creating the signals for subjective tests, types of interference signal used, methods used for decimation realization and the description of the test itself. Section 3 presents the test results. In section 4, a discussion of obtained results is given, and section 5 concludes the paper.

II. TEST SIGNALS SYNTHESIS

In this paper, A-B subjective test is used [4]. A-B test means that the subject is presented with two signals, labelled as A and B. After signal reproduction, the subject needs to answer if there was a noticeable difference between the two. Test is designed so that one of the signals (A or B) is the original signal with no processing, whereas the other signal is the one on which the processing was performed.

A. Test signals

Original test signals were taken from EBU SQAM base [6], which contains reference audio recordings for psychoacoustic tests. Sampling frequency is 48 kHz. Four recordings were selected: guitar (recording 58), speech (recording 50), music (recording 69) and violin (recording 8). These signals were selected as the representatives of various types of audio signals commonly found in real life applications.

Selected signals from EBU base differ in loudness levels, determined by ITU BS.1770-3 recommendation [7]. When performing listening tests, these loudness differences can cause changes in hearing sensitivity [8]. To avoid possible influence of this phenomenon on the test results, loudness normalization of all signals is performed. Calculated loudness levels of original signals were: -30 LUFS (guitar), -22 LUFS (speech), -26 LUFS (music) and -24 LUFS (violin). The loudness level of guitar recording was chosen as a reference. The flow chart of iterative equalization procedure is presented in Fig. 1. The initial step in the algorithm is calculating the loudness level of the signal to be normalized and determining the difference between this level and reference level value. If the absolute difference is less than a predefined threshold of 0.1 LUFS - algorithm

ends. Otherwise, level difference sign is determined and loudness level equalization is performed. If the calculated difference is positive the signal is attenuated, i.e. multiplied by a factor with values in the range (0,1]. For negative values of the difference the signal is amplified, i.e. multiplied by a factor larger than 1. In the initial iteration of the equalization process factors have values of 10 and 0.1. In subsequent iterations, the values are decreased and increased by a factor of 2, respectively. After multiplication, the signal is passed on to the next iteration and the procedure is repeated until the level difference threshold of 0.1 LUFS is reached.
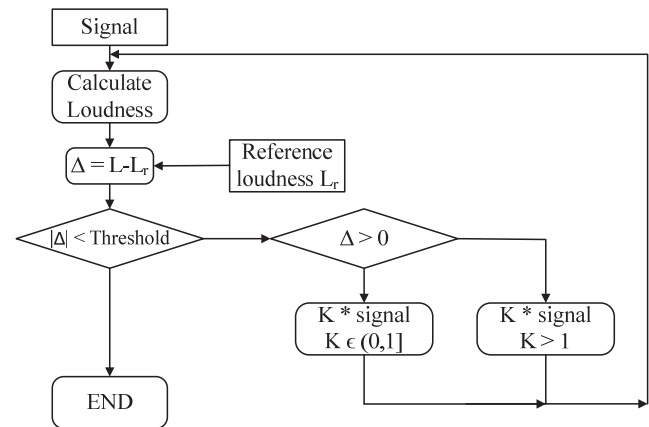

Fig. 1. Loudness level equalization algorithm.

B. High frequency noise

The original test signals do not have any spectral content above the audible range that would alias into baseband after downsampling. Therefore, an artificial noise signal with high frequency components was generated, which models the superposition of high frequency noise in audio systems. Artificial noise signal was generated with a sampling frequency of 192 kHz, normalized to 0 dBFS and then decimated to the sampling frequency of the audio signal, i.e. 48 kHz. Subsequently, the noise signal was added to original audio signal to give a final testing signal. The decimation was performed with varying decimation filter stopband attenuation, resulting in different noise interference signal levels for testing.

This procedure is artificial since the original test signal is not decimated. It should be noted that EBU SQAM audio signals are recorded with the sampling frequency of 48 kHz and that, in the process of signal preparation only the noise part of the resulting signal was decimated, i.e. filtered and downsampled. However, since the original signal does not contain any high frequency components this procedure is a good model for real life situations. Noise signal syntheses is described in [5]. Two types of interference signals were used. The first type of interference, after decimation covers nearly the entire audible frequency range, and is considered wideband. This type of interference was used in previous work [5]. The second type of interference is a signal which, after decimation, translates into frequency range of 3100-3800 Hz, which corresponds to one critical band [9]. This type of interference is obtained in the same manner as the wideband interference [5]. However, the filter cutoff frequencies and noise frequency range prior to decimation were set in such a way, so that the signal translates to the

mentioned frequency range. This type of interference is considered narrowband. Introducing the narrowband interference enables testing of the influence of interference signal frequency range on the results of subjective testing.

C. Decimation

Decimation is composed of two processes: filtering and downsampling. In this paper a decimation filter with specifications given in [1] was used. Various levels of stopband attenuations were used: 40 dB, 50 dB, 60 dB, 70 dB, 80 dB, 90 dB, 100 dB and 110 dB. Deviation of magnitude characteristic in passband is 0.1 dB. Gain characteristics of used filters are presented in Fig. 2.
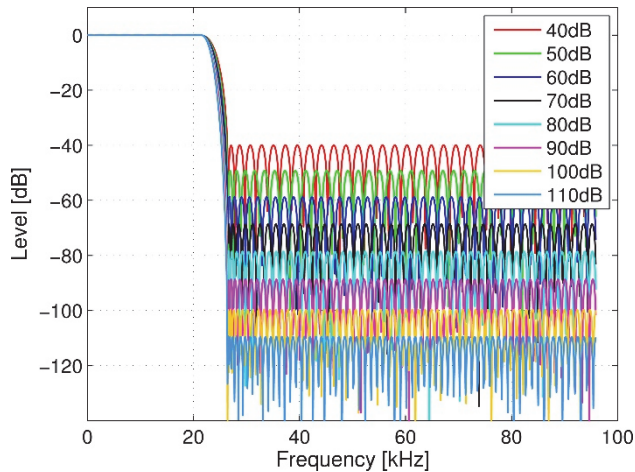

Fig. 2. Decimation filter gain characteristics.

Filters were designed as optimal filters. To satisfy amplitude characteristics for higher values of stopband attenuations, higher order filters are required. It was considered that complexity of the filter, i.e. filter order, has no influence on subjective perception. After filtering, the process of downsampling with a factor $M=4$ is performed.

D. Experimental setup

The experiment consists of two subjective tests. The signals were reproduced over headphones, which isolated the subjects from the surroundings. AKG K240 Headphones [10] and Yamaha Steinberg UR22 sound card were used [11]. For both tests, all subjects were in the age group from 20 to 30 years, for which it is considered that the sense of hearing is most sensitive. In the first test, 57 subjects participated, 18 females and 39 males. In this experiment, the wideband interference signal was used. The second experiment involved 56 subjects, 23 females and 33 males. The narrowband interference signal was used in the second experiment. In total, 70 persons participated in both tests. This means that a certain number of subjects participated in both tests. However, the testing was performed with a time interval of several days in between. To simplify the testing procedure, software with graphical user interface was written for signal reproduction and archiving of subjects' answers. The software enables reproduction of signals with all combinations of decimation filter attenuations and chosen EBU recordings - a total of 48 combinations. Reproduction of A-B pairs is random.

Random reproduction ensures that the order of A-B pairs will not influence test results. As another level of randomness, the original and the processed signal appear randomly as A or B, ensuring that the subject cannot expect to hear a difference in a certain position. Aside from mentioned 48 combinations, test includes four control pairs of A-B signals. These pairs are composed of 2 instances of the same original signal, indicating a "placebo" effect in subjects that detect a difference between them.

III. EXPERIMENTAL RESULTS

This section presents the results of subjective tests obtained for all recordings and different values of decimation filter stopband attenuation, starting with 40 dB attenuation, up to 110 dB, including equivalent signal pairs. The results for subjective tests, for two types of interference, are also presented.

Fig. 3 shows the cumulative function of dependence of noticed difference in signals from stopband attenuation values for cases of wideband interference (curve 1) and narrowband interference (curve 2). Cumulative presentation incorporates all types of recordings for a certain value of attenuation and presents a percentage of answers where the difference is detected between signals. Red dotted line shows the values of detected differences between equivalent A-B signals for the case with narrowband interference. The percentage of detected "placebo" signals in both experiments is approximately the same and is around 8%, so Fig. 3 shows only one curve.
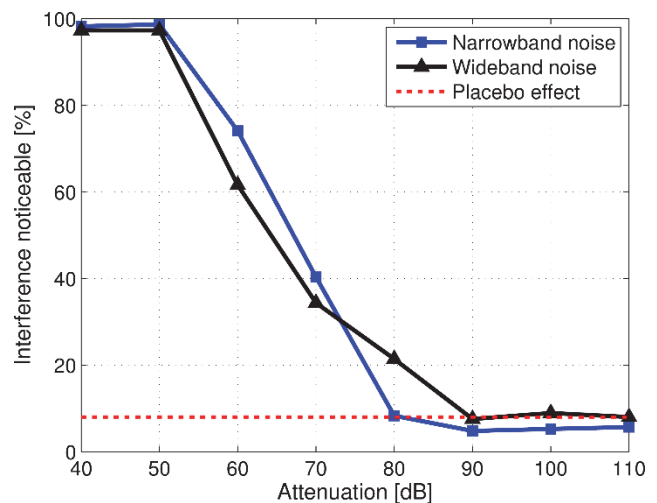

Fig. 3 Cumulative representation of noticed differences.

Fig. 4. a) to h) presents the results for different values of attenuation for the wideband interference case. Blue color signifies the number of subjects that noticed a difference between A and B signals, whereas red color signifies the number of subjects that did not detect the difference. The number of subjects is 57, so the maximal value on y-axis is 57.

Fig 5. a) to h) shows the results for different values of stopband attenuation for the case of the narrowband interference. The number of subjects is 56, so the maximum value on y axis is 56.
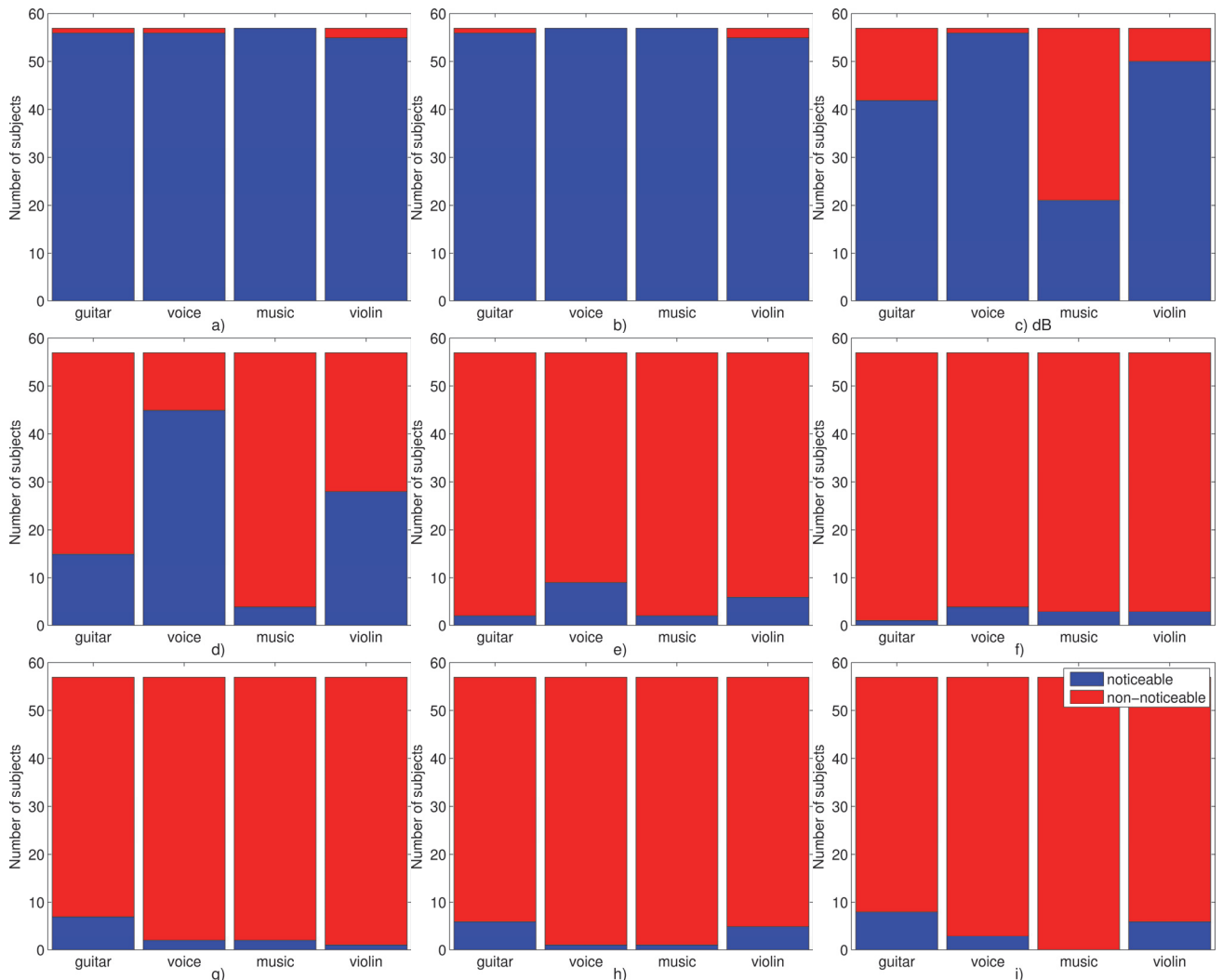
Fig. 4. Results of psychoacoustic test for different stopband attenuations and wideband interference: a) 40 dB, b) 50 dB, c) 60 dB, d) 70 dB, e) 80 dB, f) 90 dB, g) 100 dB, h) 110 dB, i) "Placebo".

IV. DISCUSSION OF THE TEST RESULTS

Experimental results were obtained based on two psychoacoustic A-B tests with 57 and 56 subjects, respectively. This number is about two times larger than in the previous experiment [5].

Noticed difference has the largest values for speech signals with both types of interfering signal. Speech signal has a narrower frequency range compared to other test signals, and the high frequency interference is aliased into certain bands of hearing frequency range where there are no components of the original signal, thereby making the difference easily perceivable. Furthermore, speech signals have periods of silence in between words, making interference even more noticeable. It can be noticed, that the detection rate of the difference is the smallest for the musical signals, for both types of interference. This fact is a consequence of a complex and wider spectral content of musical signals. Therefore, the interference is masked better than with other signals.

In this paper the interference noise prior to decimation was normalized to 0 dBFS, i.e. it has a maximal possible theoretical value. In real audio systems the peak values of interference are significantly lower, meaning that the experiment was performed for the worst-case scenario in the context of interference level.

Fig. 4. presents the cumulative representation of detected differences in the function of stopband attenuation values, as well as the value of detected differences in "placebo" tests when a wideband interference is used. The function is monotonically decreasing for both experiments. In the first experiment for the smallest value of stopband attenuation, the percentage of noticed difference is 98%. For the stopband attenuation value of 80 dB this percentage drops steeply to 8%. In the range of 80 dB to 110 dB the value of detected difference percentage is approximately constant and is around 5%. For A-B pairs of equivalent signals the percentage of noticed differences is 7%. For the second experiment, the false detection rate is almost the same, and is around 8%. The value of the stopband attenuation for which the cumulative curves cross the "placebo" floor, can be considered as a merit of the required stopband attenuation. From the results obtained from the second experiment, with the narrowband interference, it can be concluded that the required stopband attenuation is increased for about 10 dB compared to the wideband case. Due to psychological effects, there will always be a number of subjects detecting some differences. Those answers do

not necessarily signify a real difference between the signals. The overall general conclusion, which could be drawn, is that there is no need for stopband attenuation greater than 100 dB, for example. By increasing the value of stopband attenuation, the percentage of detected differences will be under 5%, and the exact value will depend on selected signals as well as on the content of high frequency noise. Aside from the fact that those high values of stopband attenuation do not guarantee complete elimination of aliased components, these filters are complicated to realize. Usually, special techniques are implemented that enables filter realization with large attenuations and steep transition regions.

Comparing the results obtained in two separate experiments, it can be concluded that the perceptibility of a high frequency interference depends also on the frequency band of that interference. For smaller values of decimation filter stopband attenuation, for both types of interference, nearly all of the subjects detected the presence of the high frequency interference. Largest differences obtained in the results for two experiments are in the case of attenuation of around 80 dB. In the case of the wideband interference the percentage of noticed differences is around 8%, whereas in the case of the narrowband interference the percentage is around 20%.

V. Conclusions

Obtained results can be useful in audio system design. The required value of stopband attenuation can be chosen based on the spectral content of the noises that can be expected in the system. It should also be noted that the value of amplitude of the high frequency signal prior to decimation was 0 dBFS in the performed experiments, which represents the worst-case scenario. In reality, the amplitudes of such interferences are substantially lower, so it can be expected that the percentage of people hearing the interference is even smaller from the values obtained in these experiments.

Literature

[1]   Lj. Milić, Multirate Filtering For Digital Signal Processing, New York, Hershey, 2009, ch.III.
[2]   B. Baker, "How delta-sigma ADCs work, Part 1 & 2", Analog Applications Journal, 3Q 2011.
[3]   M. Bjelić, M. Stanojević; J. D. Ćertić, M. Merkle "Statistical properties of quantisation noise in analogue-to-digital converter with oversampling and decimation", IET Circuits, Devices & Systems, 2017, 11, (5), pp. 421-427.
[4]   T. Poulsen, Psychoacoustic measuring methods, Oersted DTU Acoustic Tehnology, Lyngby, 2005.
[5]   M. Stanojević, M. Bjelić, J. Ćertić, D. Šumarac Pavlović, "Psihoakustička analiza uticaja parametara decimacionih filtara na kvalitet audio signala", TELFOR, Serbia, Belgrade, Nov. 2015.
[6]   EBU – TECH 3253, Sound Quality Assessment Material recordings for subjective tests, Geneva, September 2008.
[7]   Recommendation ITU-R BS.1770-3, „Algorithms to measure audio programme loudness and true-peak audio level", Geneva, August 2012.
[8]   Report ITU-R BS.2054-3, „Audio levels and loudness", Geneva, November 2013.
[9]   H. Fastl, E. Zwicker, Psychoacoustics Facts and Models, New York, Springer, 2007, ch.IV.
[10] Manufacturer technical documentation, available at: http://www.akg.com/pro/p/k240studio, accessed 2.10.2016.
[11] Manufacturer technical documentation, available at: http://download.steinberg.net/downloads_hardware/UR22/UR22_d ocumentation/UR22_OperationManual_en.pdf, accessed 2.10.2016.
[12] N. Petrović, M. Bjelić, J. Ćertić, M. Stanojević and D. Šumarac Pavlović, "Analyses of decimation filter stopband attenuation influence on subjective quality of audio signals," 2016 24th Telecommunications Forum (TELFOR), Belgrade, 2016, pp. 1-4.