Objective and Subjective Evaluation of Musical and Speech Recordings Transmitted by DAB+ System

Andrzej DOBRUCKI, Stefan BRACHMAŃSKI,
Maurycy KIN
Wrocław University of Science and Technology,
Department of Acoustics and Multimedia,Wybrzeże Wyspiańskiego 27, 50-370 Wrocław,
andrzej.dobrucki@pwr.edu.pl, stefan.brachmanski@pwr.edu.pl,
maurycy.kin@pwr.edu.pl

Abstract

The results of research on the sound quality of various kinds of music as well as speech signals transmitted via Digital Audio Broadcasting Plus system are presented. The results showed that bitrate values significantly influence the results of quality assessment, i.e. the overall audio quality as well as a timbre are dependent on the bitrate. The additional conclusion is that the CCR method is more accurate for sound assessment for higher bitrate values and this fact has been verified by standard deviation values of obtained results. The speech signals were additionally examined with PESQ method. The results have shown that the assumed quality of 4 MOS for speech could be achieved at 48 kbps. This fact was confirmed by both: subjective and objective research.

Keywords: sound quality, broadcasting, sound attributes

1. Introduction

The development of digital media technology finds the wide area of application. One of this application is Digital Audio Broadcasting (DAB) which allows not only a simple listening to the radio programs, but also some redundant-like information such as traffic, news, weather reports, city maps etc. It simply causes a growth of signal volume so it is necessary to apply some lossy compression of the transmitted signals to make all of the DAB advantages possible to work [1].

The final quality of audio signals is mainly determined by the parameters of the coding algorithms, codec quality and it is mostly independent from the parameters of transmission path in "traditional" analog meaning. It is important to evaluate how the coding as well as the transmission process influence the Quality of Service (QoS) and Quality of Experience (QoE). The first one is usually assumed as a synonymous for network performance which reflects the characteristics of devices and their parameters. However, the QoS does not reflect the real quality of signals perceived by the users, except for breaks in the signal, tracking errors etc. It was necessary to introduce a new means reflecting the impressions of listeners so the Quality of Experience can really describe the telecommunication channel from the user's point of view.

The objective quality evaluation is based on a comparison between psychoacoustic representations of tested signals and these representations are created by the mathematical

model taking into account the various aspects of hearing process including the masking phenomenon. Such psychoacoustic model defines which components or information of sound could be perceived, or not, by human ear in very particular situations. It can be found that there is a good correlation between the sound quality evaluation measured with the both: objective and subjective methods [1 - 5] although these results refer to the overall quality of sound, only, and some details of auditory image important from the aesthetical point of view have not been examined.

The main aim of the paper is to find the minimal value of bitrate of DAB+ transmission which guarantees similar hearing impressions that obtained by listening to an original signal. The results may indicate the preferences of the average listener and can give directions in the field of sound production, especially prepared for radio programs.

2. Experiment

The research on the sound quality of music and speech signals transmitted via DAB+ was performed during the emission in Wroclaw, Poland. In the experiment, the subjective quality assessment has been provided for speech and musical signals. In addition, the objective PESQ method was used for speech examination. This musical signals' assessment has been provided with standard listening tests using a Comparison Category Rating (CCR) method [6-8]. The listening tests have been provided with the staff of experts containing twelve people, aged from 26 to 31 years old. The listeners had their hearing loss no more than 5dB with the reference to normal hearing, what had been confirmed previously with audiometric tests. The listeners had participated various listening tests previously and all of them have been working as recording engineers, recording and radio producers. Tests were performed in accordance with EBU and ITU recommendations [6, 7]. The sound material contained various kinds of musical styles, from classical (chamber and symphonic) to heavy rock, and all of 13 samples were typical for the profile of National Polish Radio. The subjects of research were an overall sound quality and sound color impression. The paradigm of stimuli for musical signals was based on CCR procedure [6]: the sound samples were presented in pairs, in randomized order. The subjects were asked to rate the impairment of the second sample, in relation to the first sample, on a seven-level scale, extending from -3 (much worse) to 3 (much better). The length of sample was about 15 seconds and the musical structures of the samples were taken into account during a sample preparation. For speech signals the ACR test methodology defined in ITU [6] was used: the speech samples were rated using a scale from 1 to 5 in descriptive intervals: from excellent (5) to bad (1) according to the listeners' judgments, without comparison to the pattern. The test lists of Polish language used in the subjective speech quality measurements were created in Department of Acoustics and Multimedia, Wroclaw University of Science and Technology [5]. Each list was divided into 10 groups each with 5 tasks.

Sound samples were recorded at the analog output of the consumer set DAB Sangean DPR-26 receiver so the signal is passed by transmission channel: multiplex – DAB transmitter – receiver. The musical samples were coded with HE AAC system [1], and

transmitted with six bitrates (136 kbps, 128 kbps, 96 kbps, 80 kbps, 64 kbps and 48 kbps) with sampling frequency of 48 kHz while the speech signals were coded with five bitrates (128 kbps, 96 kbps, 64 kbps, 48 kbps and 24 kbps). As reference signals, CD recordings with the identical samples to the broadcasted ones, have been used. In addition, the same samples have been tested after passed by a simulation channel of typical FM broadcasting. The ZOOM H4n PRO was used as the recording machine.

The signals have been stored on CDR and presented with the use of CD-player (Pioneer PD-201) and a pair of active loudspeakers-boxes (TLC Pro-AMS 1). The listening sessions were provided in the recording studio of Department of Acoustics and Multimedia at Wroclaw University of Science and Technology.

3. Results and analysis

3.1. *The overall sound quality of music*

The statistical treatment by the means of ANOVA testing with the statistical power set on the critical p value of 0.05 has shown that the results could be averaged over the signals for all listeners in a group referring to particular musical style. Moreover, the variances of results obtained for the chosen groups of sound material as well as for all bitrates have been found as homogeneous and confirmed by Bartlett test ($\chi^2 < \chi_\alpha^2 = 5.99$ at $\alpha = 0.05$). Thus, it allowed to average the results over the twelve listeners in testing group and their five repetitions for four types of the musical program. The results of the overall quality evaluation and their standard deviations for 95% confidence obtained with the use of CCR method for the distinguished kinds of music are presented in Fig. 1. The results of statistical treatment have indicated a significance effect of bit-rate values ($p < 0.007$) and a musical style ($p < 0.005$).It can be seen that the quality of broadcasted musical material gets worse when the bitrate value decreases. It should also be noted that this dependence is not linear, and the signals with the lowest bitrate used in the experiment featured the worst quality, with the CMOS equal to -2. Comparing those results with the standard FM broadcasting one can assume that the similar sound quality as for FM can be achieved at 80 kbps or 96 kbps bitrates depending on the musical style.

3.2. *Sound color assessment of musical samples*

Fig. 2 presents the dependence of sound color impression obtained for different bit-rate values and the ranges of 95% confidence. Listeners' task was to evaluate the naturalness of sound color, or the change of this attribute. It can be seen that the changes of sound color get worse significantly at 48 kbps bitrate. Another interesting fact is that the sound color impression is almost the same for various kinds of musical signals for different bit-stream values, and again, this is unlike the spatial attributes of sound which featured the quality strongly dependent on types of signals [2, 3]. It means that the use of spectral band replication processing (with the HE AAC v. 1 system) makes the sound quality acceptable by the listeners even at the comparison paradigm.

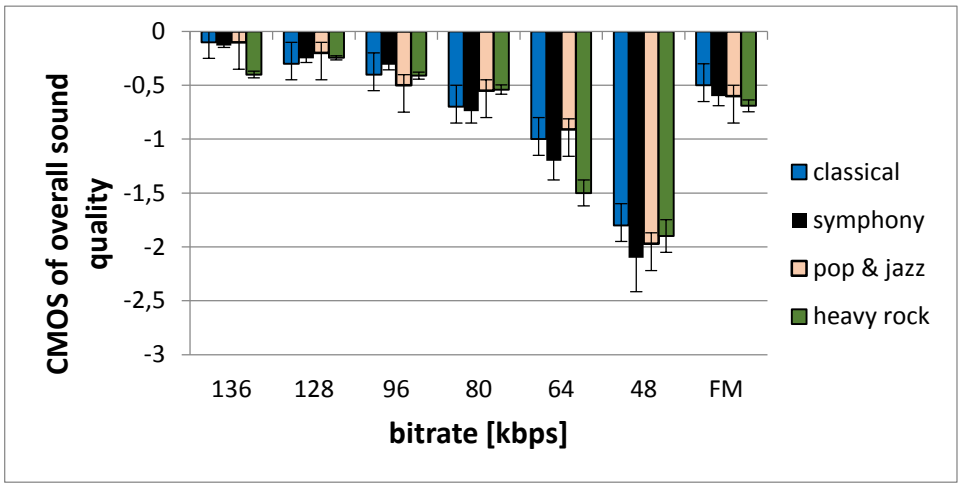

Figure 1. Results of overall sound quality obtained by the CCR method for different bitrates and various musical programs

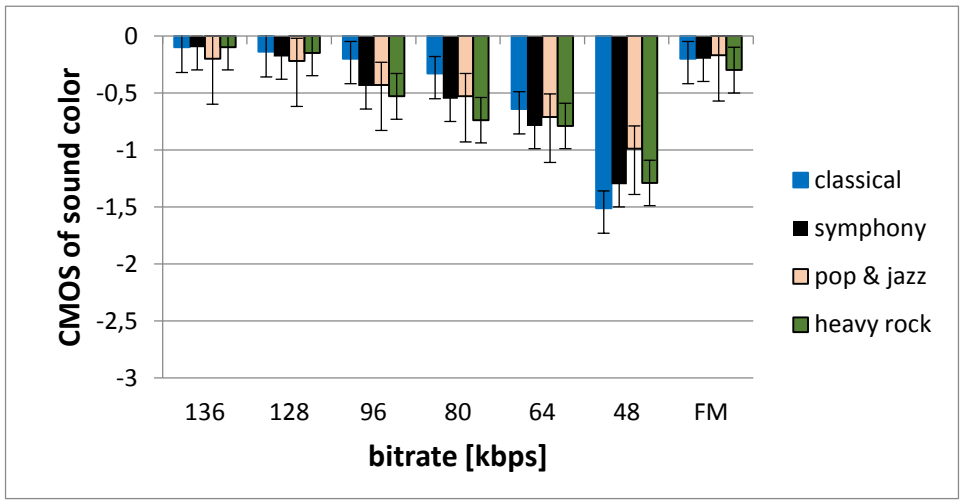

Figure 2. The dependence of sound color on the bitrate value

The results of statistical treatment indicated a significant effect of bitrate values ($p < 0.007$) only, so the significant degradation of sound quality has been observed when the bitrate values decrease. It simply suggests that for some specific groups of listeners and for specific radio program profiles, DAB+ system should be set with special parameters. For popular and jazz music when the signals are not so complex, the satisfactory sound quality is guaranteed by the bitrate of 64 kbps with SBR processing. For

more complex sound structures and aesthetical requirements (as a symphonic music, for example), the good quality may be obtained when the bit-stream is of 96 kbps, or higher.

The results obtained for FM transmission have indicated that the degradation of timbre impression is comparable to those obtained at the highest bitrates. It can also be noted that the observed changes of sound color in dependence of the bitrate have their deviation being smaller in comparison to the evaluation of the presented spatial attributes of transmitted musical signals. This fact means that the changes of spectra are more univocally interpreted by listeners from the aesthetical point of view in comparison to the spatial attributes of sound [2-4].

3.3. *Speech signal evaluation*

The testing speech material contains the lists of sentences spoken by male (M) and female (F) voices. The listeners assessed coded speech samples and hidden reference signals. The PESQ and MOS (signed as "Brachmanski") [5, 8], averaged over male and female voices, are shown in Fig. 3. This picture shows also results obtained in previous research for the MUX 1 on Gdańsk area (signed as "Gilski") [9, 10].

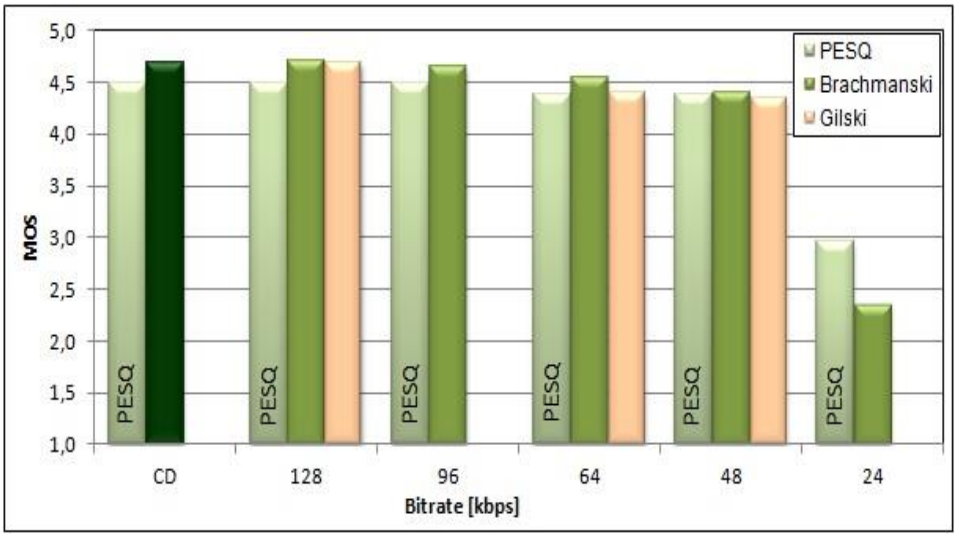

Figure 3. The MOS and PESQ results of speech signals transmitted for various bitrates

It can be shown, that for the speech signals, the minimal bit rate value which guarantees the subjective acceptable quality (4.0 MOS) has been obtained at 48 kbps bitrate. The results for this part of experiment show the good correlation between the results obtained by the objective and subjective methods of quality assessment, as this fact had been confirmed previously in the literature [3, 5]. The small shift in scores may be explained by the listeners' sensitivity for the changes in timbre of speaking voices,

but these changes are less important for the objective measurement than aspects of speech intelligibility and clarity, in a meaning of a cognitive comparison.

4. Discussion

The results of overall quality assessment of musical signals have shown a similar tendency among the bit-rate values as the notes of timbre evaluation, so it may be concluded that the good sound quality can be achieved by the use of HE AAC v.1 system at 80 kbps bitrate, if an audio material does not include a very special stereophonic tricks. It may suggest that the return of creating so called "radio version" of musical piece is recommended.

As it has been reported in literature [11] the most important factor influencing the over-all audio quality comes from the sound color impression (timbre). The other sound attributes, as spatial impression, panoramic spread, a listener envelopment etc. are less significant [2, 12, 13], although the information of all of these factors may be distorted when some processing is not transparent. However, the universally preferred values for these attributes are unlikely to exist but if the reverberation level is too low it will be easy indicated and influencing the overall sound quality [12].

The results of overall quality assessment have shown also that the greatest differences for all kinds of music have been observed for the higher bitrates. On the base of these results it could be noted that the most degradation of sound quality might occur for signals with complex time/frequency structure, i.e. symphony as well as a heavy rock pieces. This fact means that the changes of spectra are more univocally interpreted by listeners from the aesthetical point of view.

For speech signals, the results have indicated that at bitrate of 48 kbps, or higher, the MOS index takes the value of 4.5. This index is only a little bit higher than PESQ values. This fact suggest that speech quality of broadcasted voices is almost the same as the quality of original signals.

It also should be noted that the dependence of the bitrate on the subjective evaluation is not linear. It clearly means that signals with complex spectra and the time structures can be taken into account and treated as a specific material, so some tools useful for so called simple signals (as a speech, for example) cannot be applied to more complex musical structures. The reflection of this complexity in time and spectral domains could be reduced when compression of data is relatively high. During the editing process of musical recording, some effects introduced for creating certain aesthetic impressions, for example dynamic compression, are accompanied by additional processing such as equalization, stereo enhancement or reverberations. For this reason, it is still unclear whether average listeners experience the claimed negative effects of signal processing in actual recorded music productions [14]. On the other hand, signal processing with multi-layered sound effects are widely applied in a process of mastering of popular music in an attempt to boost perceived loudness. The higher degrees of compression of dynamics, for example, have been used to make musical pieces to be perceived in many different situations and environments (noisy places, stereo cars, computer loudspeakers or

mp3 players) [15] and many daily-musical-consumers do not pay attention for the lower quality of music instead of focusing on the groove an emotional aspects of the text, without reference to the pattern.

The quality of the classical FM broadcasting is almost equal to the quality of DAB+ at the bitrate value of 80 kbps, or higher. For the lower values of bitrate (48 and 64 kbps) the FM transmission offers a higher sound quality, so the additional information offered by the digital broadcasting system is the advantage of DAB+ what may suggest that these bitrates may be used for speech-based programs when particular aesthetic aspects are not important.

5. Conclusions

If one assumed that the sound quality of tested musical programs has been assessed as a little bit degraded in comparison to the original quality (with CCR value of - 0.5) the sufficient sound quality can be unconditionally guaranteed with the bitrate values of 80 kbps, or more. It can be seen that these bit-rates assure the perception of the music at the satisfactory level.

The speech signals can be transmitted with lower bitrates (the very good quality has been obtained at 48 kbps in both: objective and subjective evaluations). Moreover, the increasing the bitrate value does not make the speech quality higher significantly in a DAB+ system.

The obtained results can suggest that it is necessary to prepare the musical recording in a particular way when they have to be reproduced in proper media taking into account the technical parameters of telecommunication channel. The similar preprocessing was widely applied in classical analogue radio broadcasting (dynamic compression as well as equalization) in order to achieve almost the same sensations when listening to the radio program as to recording.

Acknowledgment

This work was supported by Wroclaw University of Science and Technology, grant no. 0401/0045/18, and National Center of Research and Development, grant no. PBS3/A3/19/2015.

References

1. W. Hoeg, T. Lauterbach, *Digital Audio Broadcasting, Principles and Applications of Digital Audio*, 2nd ed.; Wiley, England, 2003.
2. M. Kin, *Subjective evaluation of sound quality of musical recordings transmitted via DAB+ system.* Proc. of the 134th AES Convention, Rome, Italy, 2013, preprint 8874.
3. S. Brachmanski, M. J. Kin, *Quality evaluation of sound broadcasted via DAB+ system based on a single frequency network.* Proc. of the 144th AES Convention, Milano, Italy, 2018, preprint 1004.

4. A. Dobrucki, M. Kin, *Subjective and objective evaluation of sound quality of radio programs transmitted via Digital Audio Broadcast (DAB+) system.* Proc. of the ICA, Montreal, Canada, **19** (2013), doi.org/10.1121/14799170.
5. S. Brachmanski, *Quality evaluation of speech AAC and HE-AAC coding*. Joint Conference - Acoustics Ustka 2018, IEEE, 2018, 30 – 33.
6. ITU-T: Recom, P.800, *Method for subjective determination of transmission quality*. Geneva, Switzerland.
7. EBU Technical Recommendation R22-1999, *Listening Conditions for the Assessment of Sound Programme Material.* EBU Geneva, Switzerland.
8. S. Brachmanski, *Selected problems of speech transmission quality assessment* (in Polish), Oficyna Wyd. Politechniki Wrocławskiej, Wrocław 2015.
9. Gilski P., *DAB vs DAB+ Radio Broadcasting: a Subjective Comparative Study*, Archives of Acoustics, **42**(4) (2017) 715 – 723.
10. P. Gilski, J. Stefański, *Subjective and Objective Comparative Study of DAB+ Broadcast System*, Archives of Acoustics, **42**(1) (2017) 3 – 11.
11. D. Ko, W. Woszczyk, *Virtual acoustics for musicians: subjective evaluation of a virtual acoustic system in performance of string quartets.* J. Audio Eng. Soc., **66**(9) (2018) 712 – 723.
12. B. De Man, K. McNally, J. D. Reiss, *Perceptual evaluation and analysis of reverberation in multitrack music production.* J. Audio Eng. Soc., **65**(1/2) (2017) 108 – 116.
13. F. Zotter, M. Frank, *Efficient phantom source widening.* Arch. of Acoustics, **38** (2013) 27 – 37.
14. J. Hjortkjaer, M. Walter-Hansen, *Perceptual effects of dynamic range compression in popular music recordings.* J. Audio Eng. Soc., **62**(1/2) (2014) 37 – 41.
15. F. Rumsey, *Recording In The Light Of New Technology.* J. Audio Eng. Soc., **63**(12) (2015) 1053 – 1057.