Audio Quality Assessment of Digital

Radio Broadcasting Systems

Adam BARTYZAL

Dept. of Electromagnetic Field, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic  
bartyada@fel.cvut.cz

**Abstract.** Due to the amount of content in digital radio systems it is not possible to determine the sound quality of the source coding by subjective tests. In the last two decades, systems of objective evaluation using auditory models have been developed. With the development of more modern codecs utilizing psychoacoustic knowledge, machine methods seem to provide inaccurate results, i.e. they give worse quality scores than the human listener, especially at low bit rates. The paper focuses on two methods of evaluation: PEAQ and ViSQOL, their evaluation of different coding methods depending on the bitrate and tries to determine whether they are undervaluing quality and to what extent.

Keywords

audio quality assessment, codecs, PEAQ, PEMO-Q, ViSQOL, MUSHRA

# Introduction

Frequency spectrum is a very expensive matter in these days, and with the development of telephone data services, the bandwidth designated for radio broadcasting is decreasing. In contrast, the amount of content offered is growing. The solution to combine these two trends is digital radio DRM or DAB. It is necessary to find the optimal level of compression of the source signal so that the listener does not realize there is a difference from the uncompressed original and at the same time the content provider uses the spectrum as effectively as possible. To determine such a threshold, subjective tests might be used, however, their time-consumption (therefore financial cost) is high and because of that automation and computer algorithms are being introduced to provide an objective assessment.

# Objective evaluation methods

In listening tests of signals encoded by different codecs, one of the two following methods is usually used being according to ITU-R recommendations [1]. ABX (Double Blind Test) and MUSHRA (MUltiple Stimuli with Hidden Reference and Anchor). The outputs of both testing methods can be mapped to a scale defined in [1] called SDG (Subjective difference grade). Objective algorithms provide their results on the ODG (Objective difference grade) scale that corresponds to the SDG as shown in table 1.

|  |  |  |  |
| --- | --- | --- | --- |
| Impairment | Grade | SDG | ODG |
| Imperceptible  Perceptible, but not annoying  Slightly annoying  Annoying  Very annoying | 5  4  3  2  1 | 0  -1  -2  -3  -4 | 0  -1  -2  -3  -4 |

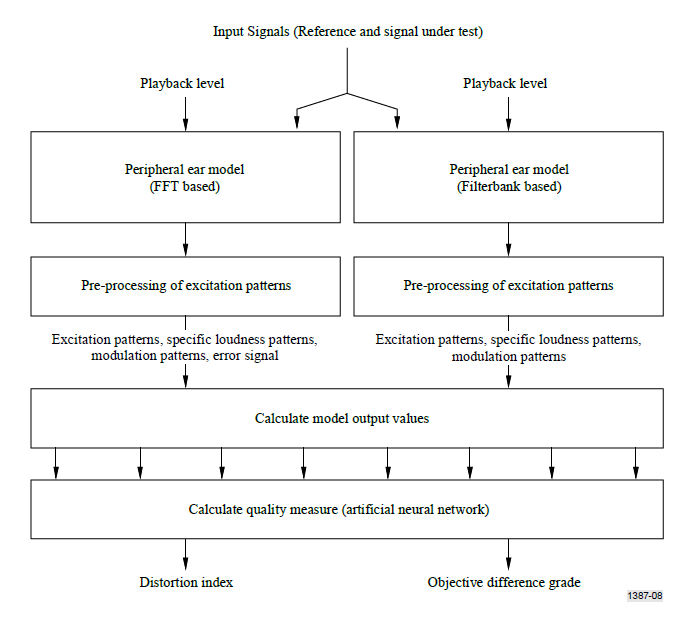
Tab. 1: Quality grades defined by [1]

## PEAQ

Perceptual Evaluation of Audio Quality defined by Recommendation ITU-R BS.1387 [2] exists in two different versions of the algorithm. Basic and advanced. There are several differences between them. The base version does not use a bank of filters to model the ear and the size of the neural network matrix used to calculate ODG.

PEAQ requires a full reference, that is, the original signal (not distorted by compression), which is compared with the test signal. Both signals pass through the psychoacoustic model, whose output are MOVs (Model Output Variables). These are then used by the artificial neural network to calculate the results as shown on Fig. 1. Detailed description of the algorithm can be found in [10].

Basic version can be obtained for free, implemented in Matlab (available at [3]). Advanced version is implemented in software called Opera [11] by a company Opticom who participated in development of the recommendation. Advanced version is being used in this article.



**Fig. 1.** Data flow in PEAQ Algorithm [10]

## ViSQOL

Unlike the previous method, an algorithm called ViSQOL (Virtual Speech Quality Objective Listener) does not try to determine the amount of distortion, noise, etc. Originally the algorithm was used for assessment of speech quality in telecommunications. Later some adjustments were presented in [1] in order to evaluate quailty of music signals. On Figure 2 is shown how the algorithm operates.

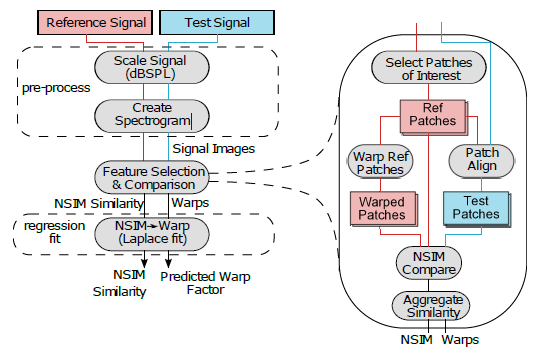


Fig. 2: Block diagram of ViSQOL Algorithm [1]

According to [4] similary to PEAQ, ViSQOL compares the distorted signal with the original, but it does it differently. In the pre-processing phase both signals are being scaled to same level. Second part of pre-processing is calculation of spectrograms using Short Time Fourier Transform STFT. Thirty logarithmically distributed filters from 250 Hz to 8 kHz simulating basilar membrane are used when assessment of speech is chosen. When evaluating musical samples, the whole hearing band is covered.

Second part of the algorithm called *Feature Selection and Comparison* takes care of selecting so-called “Patches of Interests”. Patch is a part of spectrogram with length of thirty frames. In speech evaluation three patches in bands 250 Hz, 450 Hz and 750 Hz with the highest intensity are being chosen for Mean Opinion Score computation. ViSQOLAudio overcomes this selection and uses all the patches.

Lastly spectrograms are converted to neurograms (time-neural firing activity dependencies). They are stored in form of images and then compared using NSIM (Neurogram Similarity Index Measure).

## Reference files

Because testing should reflect the actual radio broadcast, samples representing different musical genres and spoken words were selected.

At the same time, these samples were adapted for subjective testing, i.e. they were rimmed to ten seconds in duration and provided with one second fade in and fade out. List of audio samples is shown in Tab. 2.

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Filename** | **Description** | **Type** |
| 1 | capriccio.wav | Capriccio Italien Op. 45 | music |
| 2 | cimrman.wav | Part of Czech play “Opeřený Had” | speech |
| 3 | dubstep.wav | Electronic music | music |
| 4 | holmes.wav | Reading from a book: Sherlock Holmes | speech |
| 5 | pennylane.wav | The Beatles: Penny Lane | music |
| 6 | rickroll.wav | Rick Astley: Never Gonna Give You Up | music |

Tab. 2 Set of audio files

## Used Audio Codecs and bitrates

Coding of previously mentioned audio samples was prepared to standard used in DAB/DAB+ [5]. According to (mpeg) MPEG 1 Layer II can be coded into fourteen possible bitrates from 8 kbps to 384 kbps. Encoder TwoLame [6] was used set to stereo output.

Second coding method used was Advanced Audio Coding in three different profiles. First LC-AAC (Low complexity) is based on conversion to frequency domain using Discrete Cosine Transform similarly to JPEG. HE-AAC v1 adds to LC a process called Spectral Band Replication. Since last one or two octaves of audible spectrum don’t carry an important information but when they are simply cut off by low pass filter listeners describe sound as “dull”. SBR replicates lower part of spectrum to simulate higher harmonics. The third used profile was HE-AAC v2 which adds the usage of Parametric Stereo to previous profiles, combining stereo channels into one mono channel. The Spatial information is coded into three parameters Inter-channel Intensity Difference (IID), Inter-channel Cross-Correlation (ICC) and Inter-channel Phase Difference (IPD). Samples were encoded with FFMpeg [7] in following bitrates. For LC profile from 8 kbps to 256 kbps with step 4 kbps. For v2 profile the same only the upper limit is limited to 128 kbps and for v2 profile to 64 kbps.

# Results

Fig.1 shows a comparison of two selected methods. In the first line, the ODG averages of the speech samples are plotted and the second is the average of the music samples. The first column shows the PEAQ rating. For better comparison of individual codecs, the dependencies are plotted on a single graph. The second column shows the results of the ViSQOL algorithm. Its standard output is MOS-LQO, which corresponds to ratings on a scale of one to five. In comparison, however, MOS-LQO is remapped to ODG. The last column shows the difference between the two methods. It is obvious that the biggest differences are at low bitrates, which is already suggested by articles [8] and [9]. ViSQOL gives higher quality to more modern codecs.



Fig. 2: Left: Average ODG evaluated with PEAQ, Center average ODG evaluated with ViSQOL, Right: Difference between them

# Conclusion

As shown in the third column, the difference in evaluation of the two presented methods reaches up to two degrees on the ODG scale. Which one of the methods represents reality more precisely must be judged by subjective methods. The MUSHRA [2] was selected to verify this credibility. An interface for the test was prepared in Matlab (shown on Fig. 2), and preliminary results suggest that ViSQOL is closer to reality. However, it is necessary to perform the tests on a larger group of subjects, which the author will devote to in his diploma thesis.

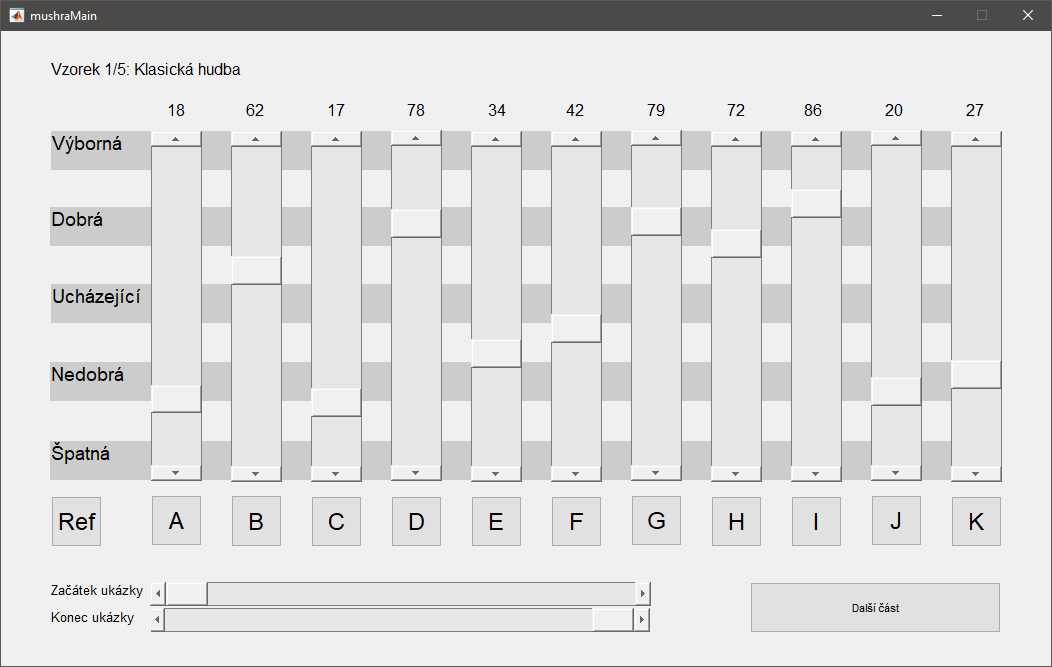


Fig. 3: GUI of MUSHRA test

Acknowledgements

Research described in the paper was supervised by Ing. K. Ulovec Ph.D., FEE CTU in Prague and supported by the Grant Agency of the Czech Technical University in Prague, under grant No. SGS17/190/OHK3/3T/13.

References

|  |  |
| --- | --- |
| [1] | ITU-R, „Methods for the subjective assessment of small impairments in audio systems“. Recommendation ITU-R BS.1116-1, 1997. |
| [2] | ITU-R, „Method for the subjective assessment of intermediate quality level of coding systems“. Recommendation ITU-R BS.1534-1, 2003. |
| [3] | P. KABAL, „GitHub,“ 2002. [Online]. Available: https://github.com/NikolajAndersson/PEAQ. |
| [4] | J. MCNALLY, Towards improving ViSQOL, Dublin: Dublin Institute of Technology, 2017. |
| [5] | E. T. S. INSTITUE, „Digital Audio Broadcasting (DAB)“. Recommendation ETSI TS 102 563 V2.1.1, 2017. |
| [6] | M. CHENG, „TwoLame,“ [Online]. Available: http://www.twolame.org/. |
| [7] | FFMPEG PROJECT, „FFMpeg Documentation,“ 2019. [Online]. Available: https://ffmpeg.org/documentation.html. |
| [8] | F. RUND a K. ULOVEC, „Comparison of Two Objective Methods of Quality Assessment for Digital Audio Broadcasting,“ *Radioelektronika,* 2018. |
| [9] | P. POČTA a J. G. BEERENDS, „Subjective and Objective Assessment of Perceived Audio Quality of Current Digital Audio Broadcasting Systems and Web-Casting Applications,“ *IEEE TRANSACTIONS ON BROADCASTING,* sv. 16, č. 3, 2016. |
| [10] | International Telecommunication Union. Rec. ITU-R BS.1387-1: Method for objective measurements of perceived audio quality. ITU-R, 2001, 100 p |
| [11] | Opticom, „OPERA™,“ Opticom, 2019. [Online]. Available: https://www.opticom.de/products/opera.php. |

About Author...

**Adam BARTYZAL** was born in Jindřichův Hradec, Czech Republic in 1994. In 2016 he obtained his bachelor’s degree in department of Electromagnetic field on Faulty of Electrotechnics, Czech Technical University in Prague and currently is continuing in master’s degree phase in the same department.



|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |