The Measurement of Perceived Quality of Various Audio Sampling Rate and Frame Loss Rate

Xiangang Qin

Quality &User Experience, Mobile Business Group, Lenovo, P.R. China
jean.qxg@gmail.com

Abstract. In this paper, the influence of Audio Sampling Rate (ASR) and Frame Loss Rate (FLR) on perceived Quality of Experience (QoE) was studied. The result indicated that users are very sensitive to the damaged auditory quality caused by frame loss at 8 kHz and 12 kHz no matter how much it losses. The perceived damage of auditory quality caused by frame loss at 16 kHz and 24 kHz is also much lower that at 8 kHz and 12 kHz. Users even failed to perceive the negative impact of frame loss on auditory quality at 32 kHz whatever the frame loss rate is. The interaction effect indicates that users are not so sensitive to the negative impact of frame loss when the sampling rates increase to 16 kHz or higher.

Keywords: Perceived Quality of Experience, Audio Sampling Rate, Frame Loss Rate.

1 Introduction

Digital audio is the fundamental media in modern digital life. Currently almost all of the musical files preloaded in Smart Phone, Tablet and Smart TV, downloaded from internet or played online are digital audios. The advancement from analog audio to digital audio has significantly reduced the costs and improved the efficiency of distribution [1].

In digital audio system, sound is passed through an analog-to-digital converter (ADC) that converts an analog signal to a digital signal. The ADC runs at a specified sampling rate and converts at a known bit resolution [2].

The sampling rate defines the number of samples per second taken from a continuous (analog) signal to make a discrete (digital) signal [3]. The range of hearing for a healthy young person is 20 to 20,000 hertz [4]. According to Nyquist–Shannon sampling theorem, perfect reconstruction of a signal is possible when the sampling rate is greater than twice the maximum rate of the signal being sampled, or equivalently. The 44.1 kHz sampling rate used for Compact Disc was chosen for this and other technical reasons [3].

D. Harris (Ed.): EPCE 2014, LNAI 8532, pp. 265–271, 2014.

Although having a sampling frequency more than twice the desired system bandwidth is desirable in some cases that extreme audio quality is required, lower sampling rates have the benefit of smaller data size and easier storage and transport [3]. That means the benefits of higher and lower sampling rates should be balanced in producing and designing the audios in a real and commercial system.

Audios are massively used in designing the interactions with users in Smart TV, such as the boot-up music, user manual, audio menu and controls and preloaded musical files, etc. The ideal audios should provide both good hearing experience that requires high sampling rate and good performance experience that requires small data size and quick transport.

Previous researches showed that most adults can't hear much above 16 kHz[5]. However, what will happen if frame is lost during transporting? Is 16kHz still the ideal sampling rate in balancing the data size and perceived quality of auditory experience.

This paper is dedicated to address this unanswered question.

2 Methodologies

2.1 Testing Stimuli

A audio file was original recorded in Chinese Language with Audacity(a audio editing software) with the sampling rate of 96kHz. It simulated a clip of typical dialogue "Xiaolin, Jin Wan You Kong Mei? Zan Lia Yi Qi Chi Ge Fan Bei' which means "XiaoLin, Are you free to have dinner with me this evening?" in English. The original audio file was then transformed into testing stimuli with various sampling rates ranging from 8kHz, 12kHz, 16kHz, 24kHz to 32kHz. The testing stimuli with various sampling were further transformed with various frame loss rate ranging from 0%, 1%, 3% to 5%. Totally, 20 (5 levels of sampling rates plus 4 levels of frame loss rates) testing stimuli were designed as the testing stimuli.

2.2 Testing Environment and Devices

The testing was conducted in a meeting room which simulated a typical living-room environment where the Smart TV was usually placed. The background noise is roughly about 45db. The audio files were played by a Lab-Top the speaker parameters of which are similar to that of Smart TV. The playing sound is about 78db. Users were seated in a chair which is about 1m away from the Lab-top

2.3 Participants

28 participants aged from 22 to 38 were invited to participate in the testing, half is Male and the other half is female. All of the participants have self-reported normal hearing ability.

Testing Procedure

The testing was conducted in four phases.

Warm-up Phase: The five audio files with various sampling rates were played in sequence from 8 kHz to 32 kHz for users to get the baseline of rating the perceived quality of experience.

1st Testing Phase: Mean Opinion Score (MOS) of sampling rates. In this phase, MOS is used to rate the perceived Quality of Experience of the five audio files with various sampling rates respectively in random. A rating scale from 1-5 was used to rate the perceived experience level of the speaking, in which 1 represents Bad and 5 represents Good. Each audio file with different sampling rate was played one time and users can ask the moderator to play the same audio file once again if it is needed.

2nd Testing Phase: Pair Comparison of Sampling Rates. In order to explore the possibility of differentiating the perceived experience between pair of sampling rates, Pair Comparison Method (PCM) was used in this phase. 10 pairs of audio files with various sampling rates were compared in random sequence. After each pair of audio files was played, user would orally tell which one is better in terms of the perceived quality of experience.

$$C_5^2 = \frac{5}{2 \times (5-2)} = \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1 \times 3 \times 2 \times 1} = 10$$

3rd Testing Phase: Pair Comparison of Frame Loss. In this phase, PCM was used to evaluate the influence of various frame loss on perceived quality of experience of audio files with various sampling rates. However, the comparisons were only made among various frame loss within one sampling rate and the frame loss across different sampling rate were not compared. Altogether, 30 pairs of audio files with various frame loss were compared.

$$5 \times C_4^2 = \frac{4}{2 \times (4-2)} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1 \times 2 \times 1} = 5 \times 6 = 30$$

3 Results

3.1 Mean Opinion Score of Sampling Rates

The result of MOS shows significant main effect of sampling rates, $F(4,112)=8.14, p<0.01$, The perceived QoE of 16Khz, 24Khz and 32Khz is significantly higher than that of 8khz and 12khz. No significant difference was found between 16khz, 24khz and 32khz. It indicates that perceived QoE of 16kHz reaches plateau and sampling rate which is higher than 16khz contributes little to the improvement of perceived QoE.

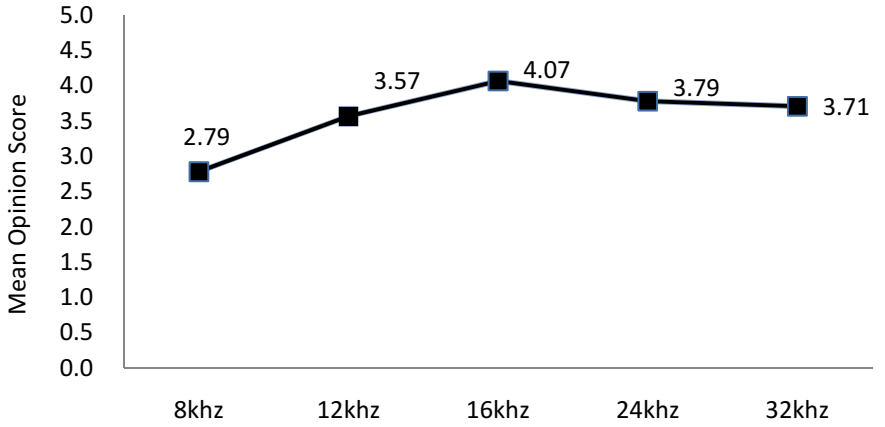

Fig. 1. The Mean Opinion Score of Various Sampling Rates

3.2 Pair Comparison of Sampling Rates

The result of Pair Comparison of Sampling Rates shows that the probability of being perceived better than a lower sampling rate declines when the sampling rate is 16kHz or higher. This result indicates that the perception of the differences between high sampling rates gets difficult.

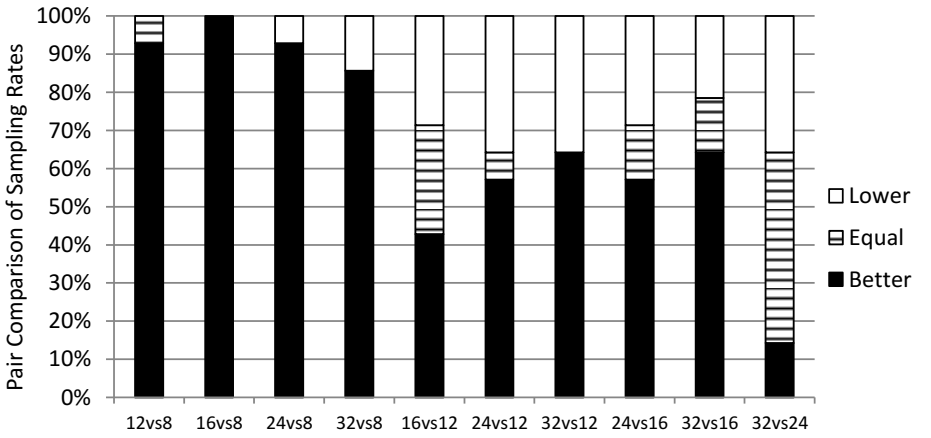

Fig. 2. Pair Comparison of Sampling Rates

3.3 Pair Comparison of Frame Loss

The result of PCM of Frame loss reveals significant main effect in the perceived QoE of six pairs (0% vs 1%, 0% vs 3%, 0% vs 5%,1% vs 3%,1% vs 5%,3% vs 5%) of various frame loss rate. *F(5,120)=38.57,p<0.01*.The perceived QoE of loss rate at 1%, 3% and 5% is significantly lower than that at 0%. The corresponding damaging value is -0.64 , -0.75 and -0.64 respectively on the -2 to +2 rating scale in which – means the perceived QoE is damaged and + means it is improved. However, no significant difference is found between loss rate at 3% and 5%.

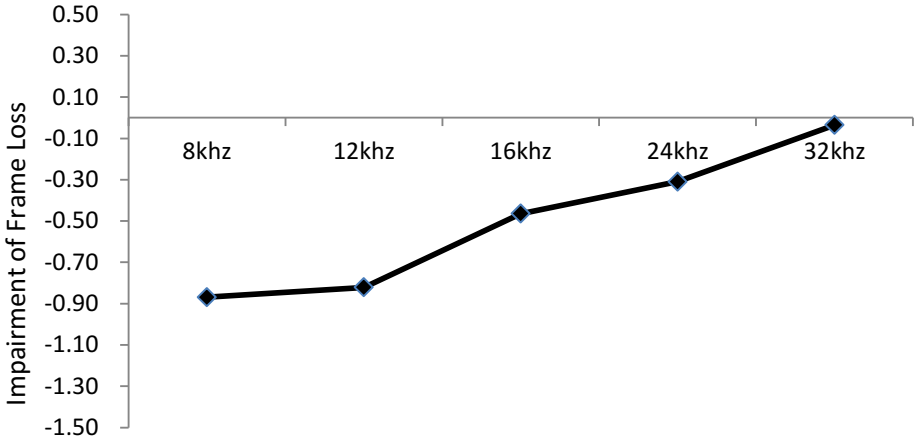

Fig. 3. Impairment of Frame Loss on Perceived Quality of Experience with Various Sampling Rates

Significant interaction effect is found in sampling rate and frame loss rate , F(20,540)=9.33,p<0.01. When the sampling rate is at 8khz and 12khz, users perceived significant damage of auditory quality at 1%, 3% and 5% frame loss rate in comparison with 0%. However, the perceived damage of auditory quality is much lower when comparison is made between various frame loss rates. The result means that users are very sensitive to the damaged auditory quality caused by frame loss at 8khz and 12khz no matter how much it losses. The perceived damage of auditory quality caused by frame loss at 16khz and 24khz is also much lower that at 8khz and 12khz. Users even failed to perceive the negative impact of frame loss on auditory quality at 32khz whatever the frame loss rate is. The interaction effect indicates that users are not so sensitive to the negative impact of frame loss when the sampling rates increase to 16khz or higher.

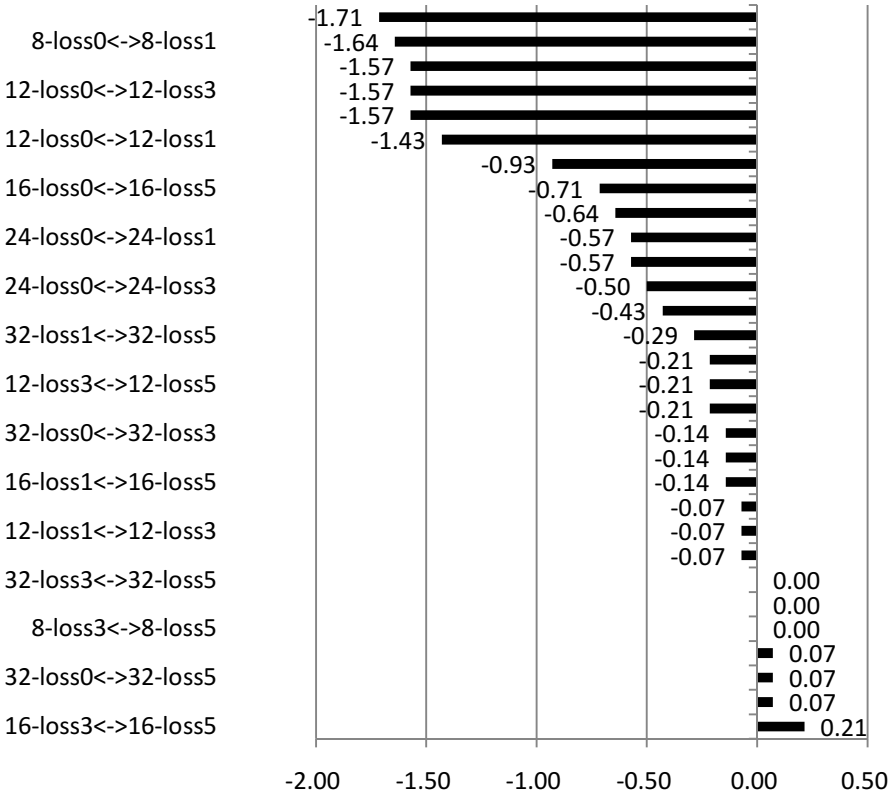

Fig. 4. The Comparison of Perceived Quality of Experience of Various Frame Loss

4 Discussion

4.1 The Measurement of Perceived Quality of Experience of Audio File

In this study, Mean Opinion Score and Pair Comparison Method were used as the evaluating methodologies. Although MOS measures the independent perception of the experience and PCM measures the dependent perception of experience which relies heavily on the relative differences between audio files, this study found consistent results in the findings by different measurement methods. It validates the reliability of this study.

4.2 16 kHz is the Turning Point of Perceived Quality of Experience of Sampling Rates

The result of MOS shows that the perceived experience increases sharply with higher sampling rates but reaches a plateau when it is 16 kHz and higher. The negative impact of frame loss on perceived QoE also decreases sharply when the sampling rate is 16 kHz or higher. This means 16 kHz is a "golden" sampling rate which makes desirable balance between perceived QoE and efficiency of data transportation.

4.3 Frame Loss Impaired the Perceived QoE Heavily with Low Sampling Rates

When audio files with low sampling rates (lower than 16 kHz) were used in designing the auditory interface of products, the insurance of data transportation is key to the satisfaction. Even only 1% frame is lost that users could clearly perceive the impairment of QoE. Audio files with 16kHz or higher is preferable.

4.4 Balance between Sampling Rate and Frame Loss

Although it's desirable to use high sampling rate in designing the auditory elements of HCI system, the redundancy of higher sampling rate shouldn't be wasted because the ability of most ordinary humans is unable to distinguish the supposed advantages of higher sampling rate over low sampling rate when it's higher than 16 kHz. The redundant resources should be used to improve the benefits of small data size, data transportation, etc.

5 Summary

Based on the results of this study, audio files with sampling rate of 16 kHz should be used in designing the auditory interaction of Smart TV which makes desirable balance between perceived QoE and data transportation.

References

1. Janssens, J., Vandaele, S., Beken, T.V.: The Music Industry on (the) Line? Surviving Music Piracy in a Digital Era. European Journal of Crime 77(96) (2009)
2. Wikipedia , http://en.wikipedia.org/wiki/Digital_audio
3. Wikipedia, http://en.wikipedia.org/wiki/Sampling_rate
4. Frequency Range of Human Hearing,
5. http://hypertextbook.com/facts/2003/ChrisDAmbrose.shtml
6. Shepherd, I.: Are high sample rates making your music sound worse?, http://productionadvice.co.uk/high-sample-rates-make-your-music-sound-worse/