Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Assessment of Audio Sound Quality Based on Psychoacoustics

Yu Ting Tsai and Jin H. Huang

Electroacoustic Graduate program, Feng Chia University, Taichung 40724, Taiwan (ROC)

PACS: 43.75.Zz, 43.75.Yy, 43.75.St, 43.75.Cd

ABSTRACT

Sound quality is an important issue in sound products today, covering a range of fields form music performance to mechanical noise, and is related to human aural response. Many measurement assessment items for sound quality have been defined including frequency and loudness. Music also includes all subjective characteristics of sound. Sound allows people to appreciate their surroundings through auditory organ, and listeners naturally anticipate enjoyment of music. In brief, “timbre” is determined by “hearing sensation” and “satisfaction” is determined by both “sound imaging” and artistic contents. When music is played, listeners pay attention to “hearing sensation” first and “satisfaction” second. But “timbre of feeling” is difficult to express objectively as listeners’ subjective feelings cannot be accurately measured by acoustic measurement equipment. To combine objective analysis and psychoacoustics to reinterpret the ratio of “timbre of feeling” in sound quality, this paper presents an assessment model for the sound quality of audio performance based on psychoacoustic theory. The model incorporates auditory roughness and specific loudness that are deemed the causes of the quality of audio performance. From the model, the optimum curve for auditory roughness is presented. Furthermore, the hearing balance of high audio fidelity and hearing satisfaction rank are proposed. Experimental results show that the model can be applied not only to measure the sound quality of audio signal but also to assess sound quality qualitative comparisons of high fidelity loudspeakers. The results also demonstrate that the proposed assessment model is capable of expressing subjective sound quality successfully.

INTRODUCTION

Sound quality is a vutal issue for every audio product that can be either subjectively or objectively. Subjective assessments measure human auditory perception directly. Given that hearing is not a purely mechanical phenomenon of wave transmission and it also involves both perceptions and cognitive processes in a person’s listening experience, subjective experiments are therefore essential even though these experiments are in general inefficient, time consuming and context dependent. On the other hand, objective methods require through investigations of the relations between the measurement and perceived quality. Theoretically this can be achieved by developing all transformation functions of human auditory system with the knowledge of human anatomy. it is desirable that subjective judgement or assessment can be replaced, or at least complemented, by an objective measurement method.

Previous studies on loudspeaker analysis, including those by Leong [1] and Hirahara [2], used frequency response measured to analyze speaker quality by machine. Kallinen [3] used frequency response measured by machine in combination with subjective hearing test to discuss sound quality. Sung [4] developed an objective measurement indicator to discuss people’s perception of timbre. Voinle and Briolle [5] subjectively discussed in sound quality in terms of mathematical transfer function. In contrast the above-mentioned studies, Fastl [6] developed a psychoacoustic theory using ear perception and psychoacoustics that to better express sound quality change. Measurement data and perception weighted curve of ear experience can be used to express some subjective characteristics. Previous studies have not accurately defined the threshold of people’s perception of sound quality. Thus, the aim of this study is to develop a model for determining such threshold.

As psychoacoustics can express ear reaction, they are often used to measure noise. Chatterley [7] compared psychoacoustic measures of sharpness, roughness, tonality and loudness used in subjective and objective testing of mechanical noise. Although psychoacoustics is often discussed in issues of noise, this does not mean that psychoacoustics cannot be applied to music. In audio compression, the most important application is Psychoacoustic-model [8, 9], using masking effect [10] to identify ear sensation and then to execute audio compression. Due to the success of MPEG coding, good and bad aspects of compression can be used to provide information on the objective indices of ear timbre as well as quality threshold in subjective timbre. Therefore, this paper established a Sound Quality Assessment Model to determine sound quality variation equilibrium to discuss the index threshold of sound quality. This model solves the problem of using hearing sensation parameters in the identification of subjective sound quality. Empirical methods included known compression signal spectrum analysis of sound quality and verification of the effectiveness of our methods. In addition, headphones were used as a measurement material to compare sound quality by music playing.

SOUND QUALITY ASSESSMENT MODEL

The sound quality assessment model uses subjective quantization methods of psychoacoustics including Auditory Roughness and Specific Loudness to observe and analyze the changes in sound spectrum. Spectrum data operation of Auditory Roughness is carried out according to the Roughness calculation methods of Vassilakis [11, 12]. The Roughness data analysis of Vassilakis uses time to express Roughness total value of each frequency. However, due to fluctuation in signal time, it is difficult to explain Roughness value of each frequency on hearing sensation. Therefore, in this paper, we divided unequal frequency filter into 24 hearing perception bands according to the methods of Bark [13]. Using Vassilakis' methods to calculate Roughness amount, Roughness amount from hearing perception band in each frequency can be clearly expressed, and this can be referred to as Auditory Roughness.

Specific Loudness calculation is carried out according to the methods of Zwicker [14], using subjective methods to represent volume loudness of each Bark frequency. If variation in loudness between high and low is small, sound balance is good. As volume loudness cannot express signal in timbre fluctuation of each frequency, it must be coordinated with Auditory Roughness described in the previous paragraph. What proportions of Specific Loudness [14-16] and Auditory Roughness can express better subjective sound? This ratio of Hearing Balance can express hearing balance to discuss sound performance. The importance of the Sound Quality Assessment Model is that the data on Hearing Balance can be used to develop summation weighted Satisfaction Rank to discuss threshold of sound quality.

Auditory Roughness Analysis Model

Figure 1 shows a flowchart of auditory roughness analysis model. Computation of the auditory roughness analysis model consists of the following steps:

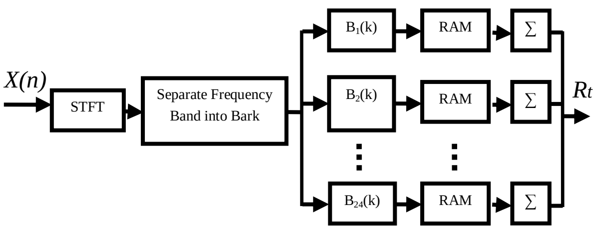

Figure 1. The flowchart of auditory roughness analysis model.

As Figure 1, According to STFT [17], Input signal $X(n)$ into a local buffer, then take a length $N=1024$ FFT of $X(n)$ to obtain the STFT at windows size m :

$$X(\omega)=\sum_{n=-N/2}^{N/2-1} X(n)W(n-m)e^{-j\omega_k n} \qquad [1]$$

where $\omega_k=\frac{2\pi k f_s}{N}$ and $f_s=\frac{1}{T}$ is the sampling rate in Hz. The STFT bin number is k . Each bin $X(\omega)$ of the STFT can be regarded as a sample of the complex signal at the output of a lowpass filter with input $X(n)W(n-m)e^{-j\omega_k n}$; this signal is $X(n)$ frequency-shifted so that frequency w_k is moved to 0 Hz. In this interpretation, the hop size R is the downsampling factor applied to each bandpass output, and the analysis window $W(n)$ is the impulse response of the anti-aliasing filter used with the downsampling. However, $X(\omega)$ is a function of two variables: time and frequency.

After obtaining time and frequency, first, cancel the weak signal below 35 dB, and then select each time with frequency to determine the filter banks as critical-band rate as in the following equation.

$$B_n=13*\arctan\left(0.76\frac{f}{kHz}\right)+3.5*\arctan\left(\frac{f}{7.5\cdot kHz}\right)^2 \qquad [2]$$

where B_n is the index number in the frequency scale. The individual critical bands have bandwidths Δf_b as follows.

$$\Delta f_b=25+75\left(1+1.4\left(\frac{f}{kHz}\right)^2\right)^{0.69} \qquad [3]$$

Eq. 2 and Eq. 3 describe the dependence of critical-band rate also called Bark [14]. It is important to implement the filter design. The bark computation is based on a recursive algorithm.

The auditory roughness level is dependent on the amplitude fluctuation and results in level dependency on every frequencies. This corresponds to the known definition of audio with different frequencies ($f_{\max}$ and $f_{\min}$) and amplitudes (a_1 a1 and a_2) as follows.

$$R=(a_1\cdot a_2)^{0.1}\cdot 0.5\cdot\left(\frac{2\min(a_1,a_2)}{a_1+a_2}\right)^{3.11}\cdot d(x) \qquad [4]$$

where

$$d(x)=\left(e^{-3.5\cdot s(f_{\max}-f_{\min})}-e^{-5.75\cdot s(f_{\max}-f_{\min})}\right)$$

and $s = \dfrac{0.24}{0.0207 f_{\min} + 18.96}$ [5]

The term $(a_1 \cdot a_2)^{0.1}$ represents the dependence of roughness on intensity of added sinusoidal signals, which is related to their amplitude. The term $\left(\dfrac{2\min(a_1, a_2)}{a_1 + a_2}\right)^{3.11}$ represents the dependence of roughness on amplitude fluctuation degree of added sinusoidal signals, which is related to their amplitude difference. Finally, the term $d(x)$ represents the dependence of roughness on the amplitude fluctuation rate, which is related to frequency difference and depends on the frequency of lower sinusoidal signal. For every roughness resolution during the temporal domain, each roughness R has to consider the unnecessary amplitude fluctuation. Furthermore, the roughness of signals corresponding to spectra with more than two sinusoidal signal components can be calculated by summing the roughness of all sinusoidal-pairs in the spectrum. It is found that, depending on the relative phase of the respective amplitude fluctuations, the total roughness may be less than the sum of the roughness values for individual pairs. Therefore, the total roughness can be summed over all auditory filters yielding as following,

$$R_t = \frac{1}{\Delta T} \sum_{t=T}^{T'} \left[\frac{1}{\Delta f_b} \sum_{n=f_L}^{f_L'} R(n) \right] \quad [6]$$

where R_t is total roughness according to every Bark scale, t is the time range $T \rightarrow T'$ of R, the frequency range Δf_b has been given in eq. 3. However, the signal may vary in the temporal domain, the total roughness value from the frequency viewpoint needs to consider non-signal period. So, the total Roughness value must be accurate over a specific time period.

Specific Loudness

Loudness is a sound level perceptual measure of the effect of energy content of sound on the individual's ear which is related to sound pressure level, but is not the same. Loudness is the attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from soft to loud. Fundamental assumption of model of loudness indicates that it is not a product of spectral lines and is not obtained from the spectral distribution of the sound directly, but the total loudness is the sum of specific loudness from each critical band. Specific loudness N' is given by

$$N'(f) = 0.0635 \cdot 10^{0.025 \cdot E_{TQ}} \cdot [11 + (\frac{1}{4} \cdot 10^{0.1 \cdot E - a_0 - E_{TQ}})^{0.25} - 1] \quad [7]$$

where E_{TQ} is the excitation at threshold in quiet, E_0 is the excitation corresponding to reference intensity (I_O = 10^{-12} W/m^2), a_0 is attenuation factor and E is the excitation level. Computer code in BASIC is provided by Zwicker [12]. Perceived sound quality of the signal and its relation to the various physical properties has been the subject of human listening comprehension.

Hearing Balance and Hearing Satisfaction Rank

Hearing balance is the ratio of the total roughness R_t and specific loudness N' as follows,

$$H = \frac{1}{N'/R_t + 1} \quad [8]$$

In the hearing of music, not only the frequency difference, but also the fluctuation in audio signal, affects the hearing perception and sound quality. Eq. 8 indicates the fidelity and loudness in human sensation. The Satisfaction Rank H_s can be obtained over 24 barks as follows.

$$H_s = a \sum_{n=1}^{24} H_n \quad [9]$$

where H_n is hearing balance for particular bark scale. Eq. 8 estimates the satisfaction rank/grade in the music signal. The coefficient a is designed to upgrade the result of Eq. 9. High and low frequency sound quality changes are reflected in Eq. 8. The treble, middle and bass are often influenced by the audio signal, hence we could use Eq. 9 to obtain the average hearing balance and to observe the Rank of sound quality of audio performance.Paragraphs immediately following their headings are to be justified on both sides with no indents for first lines. Use single line spacing throughout the entire document. There is a single line space between paragraphs.

EXPERIMENT

Five high fidelity headphones were selected judiciously. Headphones were typical studio headphones of various standard brands available in the market. Their cost variation is in the range of 100-180 US$ (Nov. 2008). We have designated them A1 to A5. A1 is circumaural and semi-open headphone. A2 is circumaural and closed-back headphone with auto-shut-off feature. A3 is circumaural and closed-back headphone. A4 is circumaural and open style headphone and A5 is circumaural and closed-back headphone. The observable audio quality effects of these headphones are used for analysis and comparison of their audio quality reproduction. The experimental measurements were carried out in an anechoic chamber by B&K electroacoustic equipment in accordance with the arrangements shown in Figure 2. We also used software SoundCheck 8.1. The hardware connections were followed by calibration of devices using a B&K amplifier (type 2716C) and a B&K standard calibrator (type 4231, 1 kHz sound at 94 dB ±0.2 dB). B&K calibrator was also used to calibrate B&K HATS simulator. A sweep sine wave signal from 20 Hz up to 20 kHz (human audible frequency range) was used as an input to the audio system. Signal from amplifier was recorded and processed with SoundCheck 8.1. Headphones were placed on B&K HATS for measurement. Stimulus signal (classical music, the song "Somewhere in Time", conducted by Jeannot Szwarc, 1980.) was used. Classical music is often distinguished by its wide use of instruments of varying tones and pitches to create a deep and rich sound. We have selected test signal in the range of 1:01 ~ 1:05s. Sound level did not change significantlty in this range.

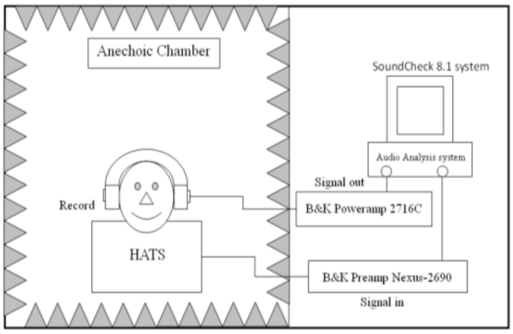

Figure 2. The schematic diagram for record headphones.

RESULTS AND DISCUSSION

The Headphone Quality Comparison

Figure 3 illustrates the auditory roughness for each headphone along with the reference signal using sound quality assessment model. Curves are similar and nearly overlapping within low to mid frequency range (up to 17 bark), but above 17 bark, the change in auditory roughness is observable. The enlargement of the above figure within the range of 16 - 24 bark is shown in Figure 4. Reference signal's auditory roughness is almost linear, however the performance of headphones varies from each other and even from the reference. It is totally specific to performance, characteristics and design. However, it is found that all auditory roughness curves converge to linear or near linear indicating a good sound quality performance.

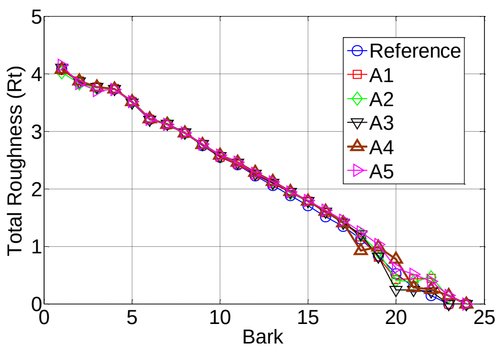

Figure 3. Auditory Roughness for headphones along with reference signal.

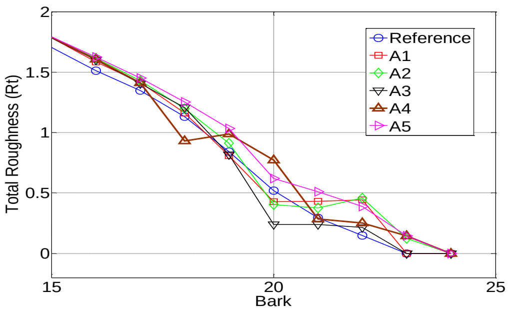

Figure 4. Auditory Roughness for headphones along with reference signal (15-24 barks).

Specific loudness curves for each headphone along with reference signal is shown in Figure 5. The overall pattern of these curves is similar, but it is still not exactly the same over each bark. This was found that the reference signal in the high frequencies (10 - 18 bark) is quite different from the other headphone signals. It may be due to consideration of effect of hearing perception while recording. The reference signal in digital form is supplied directly to analyser. However, the signals from headphones may be affected by hearing sensation due to HATS. Also, it was found that the specific loudness in the low frequency region (2 - 4 bark) for headphone A5 was lower than for other headphones, but in high-frequency region (15 - 18 bark) specific loudness was higher than for any other headphones. This indicated the unique charatcterictics of A5. The frequency response for A5 is relatively flat, hence the music reproduction would be more balanced.

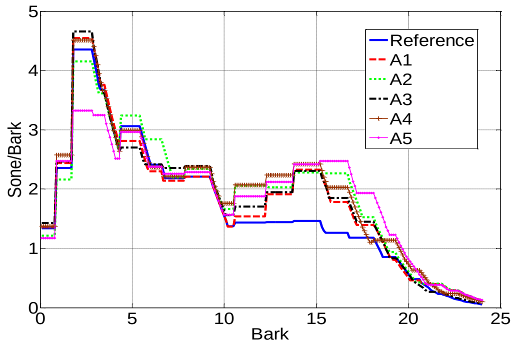

Figure 5. Specific loudness for different headphones.

Figure 6 shows the hearing balance for each headphone set with reference signal. The satisfaction rank of frequency performance with different audio compression is shown in Table 1. The $a = 7.8125$ is defined in Eq. 9 as in the previous section. Table 2 shows that the headphones A1 and A5 have higher satisfaction rank, whereas A3 has the lowest. Correspondingly, as shown in Figure 6, A1 and A5 have higher hearing balance among the five headphones within 20 – 24 bark, whereas A3 has the lowest. For the high frequency region (20 – 24 bark), A3 and A5 represent the worst and best cases. Visual inspection of the hearing balance curve (more fluctuations) for A3 confirms that it is showing/reproducing irregular sound and appears to have some rough or disordered sound. The performances of A1 and A5 seem to be much more regular. Similarly, A2 has more creaky sounds than A4 despite a higher satisfaction rank for A2. Figure 6 and Table 1 show that the reference audio has greater hearing perception when compared with the headphones.

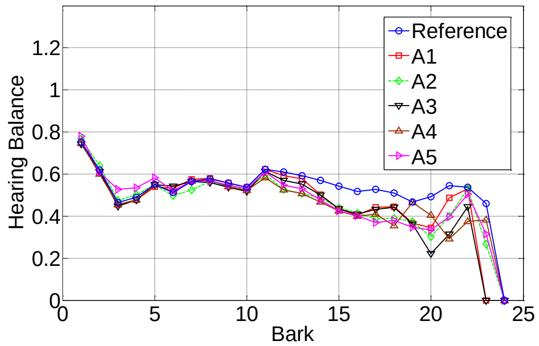

Figure 6. Hearing balance of different headphone with reference signal.

Table 1. Satisfaction rank with different headphones sound quality.

Name	Satisfaction ranks
Refence	98.5794
A1	88.7562
A2	87.5728
A3	84.6596
A4	86.9693
A5	89.0983

Audio Compression Quality Comparison

Figure 7 presents a general audio compression coding flowchart which is a simplification of MPEG Layer-III audio compression model. The input signal $X(n)$ and output signal $Y(n)$ are both mapped onto a psychophysical representation by means of three operations, frequency warping, time-frequency smearing and compression of the intensity scale. The masking effect of psychoacoustic model predicts the lossless and lossy of compression rate which are based on reconstruction of audio signal after FFT block [18, 19] and then encoded onto $Y(n)$ which is affected by different masking levels. In our tests, the audio quality is influenced by controlled individual masking threshold.

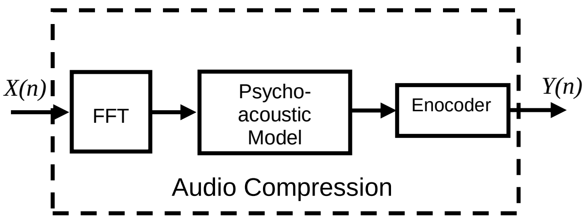

Figure 7. Flow chart of audio compression.

The psychoacoustic model predicts the final masking level on the encoder. When masking is exploited in audio bit rate reduction, the final masking level makes up a significant portion of the whole frequency-time space. Enforcing variation in the masking level enables destruction of or variety in the subjective perceptual audio quality cause by the masking threshold. The masking index for tonal masking components is given by Eq. 10 [20] with a constant *c*.

$$a_{vtm} = -2.025 - 0.275 * z(j) - c \quad [10]$$

where a_{vtm} is tonal masking components, *z(j)* is the frequency index of the bark scale in which j is the index label of tonal label. Treat the constant c as the variable for the purpose of making the wrong tonal estimation of audio quality of signal which is helpful to figure out good quality audio compression. Comparison of the output of the audio compression codec for masking and un-coded masking of reference signal is good for understanding the satisfaction level of audio performance. A good result for c is equal to 3.5 in the MPEG standard [20]. This could be used to explain the good satisfaction rank in this test.

Figure 8 shows the auditory roughness curves with different audio compressions. Instead of a constant value for *c*, its effects on auditory roughness, specific loudness and hearing balance were investigated to figure out their importance. It was found that by increasing variable *c* (higher compression ratio) the audio quality became worse. Hence it can be said that it is the wrong estimation of the final tonal masking level. Also, curves in Figure 8 present the quality of audio fluctuation in the signal. This indicates the timbre effect of the frequency range on hearing response. It was also found that if the curve becomes nonlinear, the timbre worsens. This is because the dissonance responses of hearing perception are correlated with auditory roughness. The reference curve shows the best sound quality for original audio signal and the balance of auditory roughness in every bark. For the value of *c = 5*, we observe poor quality audio reproduction.

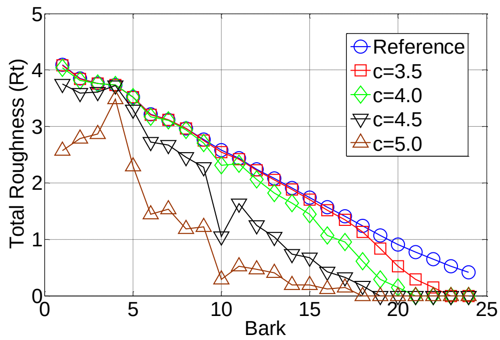

Figure 8. Auditory roughness curves with different audio compression.

The specific loudness curves in Figure 9 show the variation in specific loudness with different audio compressions. It is evident from the results that the variations in *c* do not change the specific loudness level. However, auditory roughness curve shows significant changes with these variations. Specific loudness is not fully able to interpret the best timbre for music reproduction. A very small change in the curve with *c=5.0* is observed corresponding to 15 to 22 bark (high frequency). This definitely represents the distortion of loudness due to audio compression.

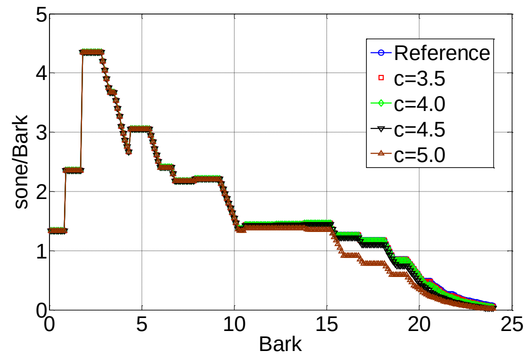

Figure 9. Specific Loudness with different audio compression.

The hearing balance curves in Figure 10 show how auditory roughness and specific loudness collectively affect the hearing balance of sound quality. It is a pertinent observation in the figure that with the increasing values in *c*, the curves depart from the reference curve at different barks. As we have already explained, the variation of specific loudness with variation in *c* is very small, hence the hearing balance is dependent on the sound fidelity only. In addition, change in audio compression ratio has the most direct impact on the mid to high-frequency signal range (5 – 24 barks).

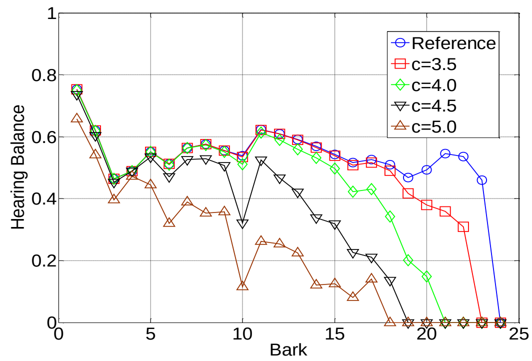

Figure 10. Hearing balance of curves with different audio compression.

The satisfaction rank of frequency performance with different audio compressions is shown in Table 2. The $a = 7.8125$ is defined in Eq. 9 of this result. As the summation of hearing balance is 12.6, we did not make the satisfaction rank higher than 100. If the satisfaction rank decreases significantly, sound reproduction quality worsens. If audio compression goes beyond the limit (77 or below), sound quality will become very bad. This clearly signifies that the sound becomes distorted and discordant.

Table 1. Satisfaction rank with different audio compression.

Name	Satisfaction ranks
Refence	98.5794
c=3.5	90.0123
c=4.0	77.4940
c=4.5	61.0607
c=5.0	41.0923

The Optimum Auditory Roughness Curve

A variety of music signals are directed to the analyzer in digital form for auditory roughness comparison. As long as the sound quality of sound source is good, the results should be similar and tend to be linear. From the analysis of four good CD-quality songs of different types in Figure 11(a), we found that results are similar and almost linear. This means that auditory roughness alone can obtain an answer as to the best sound quality reproduction irrespective of the sound signal type. For different kinds of music with good quality mix down and reproduction, by using the regression analysis, the best result for the roughness curve may be approximated by

$$y = -0.1692x + 4.426 \qquad [11]$$

where *x* ranges between 1 – 25 corresponding to bark. The result of the above equation is shown in Figure 11 (b).

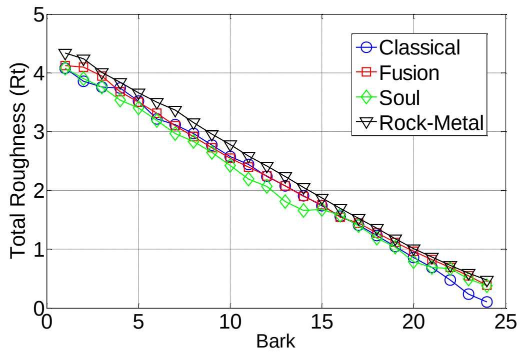

Figure 11(a). Auditory roughness curves - Four kinds of music.

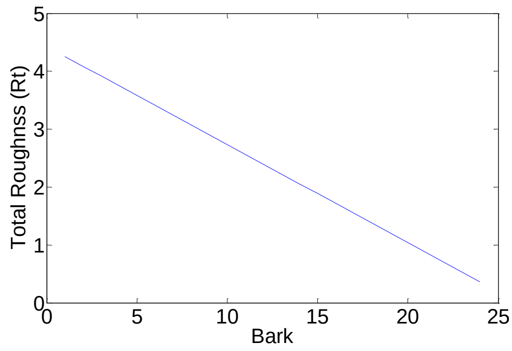

Figure 11(b). Auditory roughness curves - Optimum curve.

CONCLUSION

This paper assessed sound quality based on psychoacoustics. The results clearly show that subjective quantization methods indicators can be used to express subjective sound quality. The Auditory Roughness index can be used to determine if, in terms of frequency, is complete and whether fluctuation is excessive. In addition, Specific Loudness can subjectively describe if people's reaction to loudness is balanced at all frequencies. It can sufficiently reflect opinions regarding subjective response to loudness. The ratio present the integrated response level of these two auditory characteristics of Auditory Roughness to Specific Loudness may also observe the variation of sound quality. In our experiments, we compared ratios of audio compression. Compression rates associated with good wound quality are known due to defined constant compression ratio in subjective aspect. Thus, we can clearly observe the level of completeness of sound reproduction complete and the level of satisfaction of auditory response. Using headphones as measurement material to determine defects in sound reproduction fidelity and to implement improvements has greatly contributed to the design and quality management of headphones. Sound signal evaluation methods have greatly contributed to academic research and commercial development of speaker. The Sound Quality Model can be used not only in evaluation of industry but also in evaluation of timbre music performances.

In this study, although good and bad timbre responses are determined theoretically, quantitative differences cannot be ascertained in music preferences among different ethnicities. In future work, Comparisons of subjective testing and results are suggested with Jury Test, using a variety of music and different ethnic groups to obtain better Hear Balance hearing sensation curve.

ACKNOWLEDGEMENT

The authors would like to thank the National Science Council of Taiwan for financially supporting this research under contracts Nos. NSC 95-2221-E-035-013-MY3, 97-2221-E-035-031-MY3, and 96-2221-E-011-MY3. We also thank Listen Inc. for supplying SoundCheck measurement and Merry Electronics CO., LTD. for supplying headphones.

REFERENCES

1 P. H. W. Leong, Y. S. Moon, W. K. Sim, and D. W. P. Lam, "Sound Quality Measurements in Headphones", *106th Audio Engineering Society (AES) Convention Munich, Germany*, 4874 (B6), pp. 1 - 12, 8 - 11 May 1999.
2 T. Hirahara, "Physical Characteristics of Headphones used in Psychophysical Experiments", *Acoust. Sci. & Tech.*, Vol. 25, No. 4, pp 276 - 285, 2004.
3 K. K. N. Ravaja, "Comparing Speakers Versus Headphones in Listening to News from a Computer – Individual Differences and Psychophysiological Responses", *Comput. Hum. Behav.*, Vol. 23, Issue 1, pp. 303 – 317, Jan. 2007
4 H. Y. Sung, J. Kim, S. Jang, "A Method for Objective Sound Quality Evaluation of Headphones", *AES 32nd International Conference, Hillerød, Denmark*, pp. 83 - 87, 21–23 September 2007.
5 T. Voinle and F. Briolle, "Transfer Function and Subjective Quality of Headphones: Part 1, Transfer Function Measurements," *11th International AES Conference, Audio Test & Measurement, Portland, Oregon*, pp 248 - 253, 29-31 May 1992.
6 H. Fastl, "The Psychoacoustics of Sound-Quality Evaluation", *Acustica*, Vol. 83, No. 5, Sept./Oct. 1997.
7 J. J. Chatterley, "Sound Quality Analysis of Sewing Machines", *Master of Science Thesis, Brigham Young University, Department of Mechanical Engineering, Provo, UT, USA*, August 2005.
8 U. Zölzer, *Digital Audio Signal Processing, 2nd Edition*, A John Wiley and Sons Ltd, Publication, 2008.
9 A. J. S. Ferreira, "Perceptual Audio Coding and the Choice of an Analysis/Synthesis Filter Bank and Psychoacoustic Model," *104th Audio Engineering Society (AES) Convention Amsterdam*, 4671 (P9-5), pp. 1 - 36, 16 - 19 May1998.
10 J. G. Beerends, "The Role of Informational Masking and Perceptual Streaming in the Measurement of Music Codec Quality", *AES the 100th Convention*, 1996 May 11-14.
11 Vassilakis P. N., "SRA: A Web-based Research Tool for Spectral and Roughness Analysis of Sound Signals", *Proceedings of, 4th Sound and Music Computing Conference (SMC'07), Lefkada, Greece*, pp. 319 - 325, 11 - 13 July 2007.
12 Vassilakis P.N., "SRA: An online tool for Spectral and Roughness Analysis of sound signals", *J. Acoust. Soc. Am.*, 120(5/2): 3677.
13 E. Zwicker and E. Terhardt, "Analytical Expressions for Critical-Band Rate and Critical Bandwidth as a Function of Frequency", *J. Audio Eng. Soc.*, Vol. 68, No. 5, pp. 1523 - 1525, Nov. 1980.
14 E. Zwicker, H. Fastl, U. Widmann, K. Kurakata, S. Kuwano and S. Namba, "Program for Calculating Loudness According to DIN45631 (ISO 532B)" *J. of Acoust. Soc. of Jpan. (E)*, Vol. 12, No. 1, pp. 39 - 42, 1991.
15 J. E. Appell, "Loudness Models for rehabilitative Audiology", *Doctor of Science dissertation, University of Oldenburg, Germany*, February 2002.
16 R. M. Aarts, "Calculation of Loudness of Loudspeaker during Listening Tests", *J. Audio Eng. Soc.*, Vol. 39, No. 1/2, pp. 27 - 38, Jan./Feb. 1991.
17 J. O. Smith and X. Serra, "PARSHL: An Analysis/Synthesis Program for Non-Harmonic Sounds Based on a Sinusoidal Representation," *Proceedings of the 1987 International Computer Music Conference, International Computer Music Association, San Francisco*, pp. 290 - 297.
18 C. M. M. Tio, R. K. C. Tan and C. K. Yeo, "Psychoacoustic Parametric Audio Coding Technique Using a High-Order QM Sub-Band Filter Bank", *107th Audio Engineering Society (AES) Convention, New York*, 5046 (N-2), pp. 1 - 18, 24 - 27, September 1999.
19 R. Sušnik, J. Sodnik, S Tomažič, "An Elevation Coding Method for Auditory Displays", *Appl Acoust.*, Vol. 69, Issue 3, pp. 233 - 241, March 2008.
20 ISO/IEC 11172-3, "Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 3: Audio", *International Standards for Business, Government and Society*, 1993.