

Normalized Cumulative Spectral Distribution in Music

Young-Hwan Song, Hyung-Jun Kwon, and Myung-Jin Bae

Abstract—As the remedy used music becomes active and meditation effect through the music is verified, people take a growing interest about psychological balance or remedy given by music. From traditional studies, it is verified that the music of which spectral envelop varies approximately as $1/f$ (f is frequency) down to a frequency of low frequency bandwidth gives psychological balance.

In this paper, we researched signal properties of music which gives psychological balance. In order to find this, we derived the property from voice. Music composed by voice shows large value in NCSD. We confirmed the degree of deference between music by curvature of normalized cumulative spectral distribution. In the music that gives psychological balance, the curvature shows high value, otherwise, the curvature shows low value.

Keywords—Cognitive Psychology, Normalized Cumulative Spectral Distribution, Curvature.

I. INTRODUCTION

IN automatic recognition and artificial intelligence, the importance of emotional recognition is growing up because mechanism and algorithms are operated like human with emotional thinking. The methods to extract emotional information from signal and transfer through signal have been presented and developed. In speech signal, emotional recognition is performed by the difference of factors which is used for speech recognition through emotional changes. The methods to visualize these emotional properties of signal have also been developed and used.

For music recognition, it is classified into genre classification [1], instrument classification [2], music search [3] etc and it is performed by 90 factors approximately. Many of these factors are from speech recognition and audio description of MPEG.

It is medically proved that music effects on the psychological and mental health. So music is recently used for therapy and mediation [4]. Even music is developed and produced for these purposes [5]. Analysis of music which is used for therapy and mediation is performed by brain wave and psychological changes from the effects on the brain and body of human [6]. The music which is used for therapy or medical purpose is produced by using the similarity from the result of medical

experiments [6].

Even though the effects of emotional factor of music on life and psychological health are proved, it is not enough to express and classify the music for emotional recognition by signal own properties since the factors which can represent emotional parts are not enough.

In this paper, we researched how can express emotional elements in music objectively. So we proposed curvature of normalized cumulative spectral distribution as new factor that can represent signal property of music which gives psychological balance and improve emotional recognition rate of music.

II. MUSIC COMPOSED BY VOICE

Music is growing up with musical instruments. Some kinds of music have only played with instruments, and it also played by voice with instruments or not. Richard F. Voss and John Clarke researched that the music which is played for a long time is characterized as $1/f$ on low frequency. And this is revealed is speech [7, 8]. Namely, popular music has similar property to speech in low frequency. For music, Fourier analysis has used before. And the properties from music in frequency are applied into content-based music classification and recognition. Now we analyze the music only composed by voice. It is not just speech. After that it is compared with other music to know which music has the similar property.

A. Normalized Cumulative Spectral Distribution

In this paper, we propose Normalized Cumulative Spectral Distribution (NCSD) as factor which can express spectral distribution of music to compare with music only composed by voice. NCSD is defined by (1).

$$S_{CN}(f) = \int_0^f \frac{\log(S_{scan}(r)) - \min(\log(S_{scan}(r)))}{\int_0^{f_s/2} \log(S_{scan}(k)) - \min(\log(S_{scan}(k))) dk} dr \quad (1)$$

With NCSD, we can confirm the relative spectral distribution. It also expresses cumulative energy with increasing frequency. In speech generation model, excitation convolved with vocal track which means generation system. So spectral envelope followed around formant envelop in frequency. In formant of speech, generally energy is concentrated in low frequency. And it is decreased with increasing frequency. On the base of this, NCSD of speech is rapidly increased on low frequency, and an increase is

Manuscript received June 1, 2008.

Y.H. Song, H.J. Kwon, and M.J. Bae are with Information and Telecommunication Engineering Department, Soongsil University, 1-1 Sangdo 5 dong DongJak-Ku, Seoul, 156-743, Republic of Korea (phone: +82-02-824-0906; fax: +82-02-824-0906; e-mail: song0one@yahoo.co.kr, mswg1@naver.com, mjbae@ssu.ac.kr).

decreased with increasing frequency.

The following figure shows spectrum and NCS D of music composed by voice. In this paper, we would like to know property of popular music which makes people comfortable for a long time. So we used sacred song for this property.

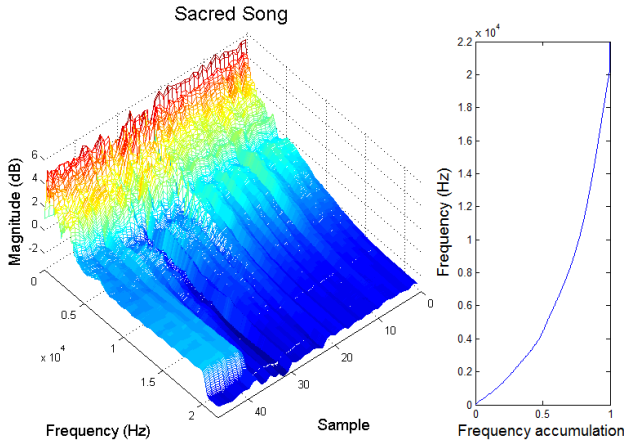


Fig. 1 Spectrum and NCS D of Sacred Song

B. NCS D Curvature

In order to represent the NCS D by one factor, we defined NCS D curvature as (2).

$$k_{CN} = \frac{L_{CN_envelop}}{L_{straight}} \quad (2)$$

Where $L_{CN_envelop}$ is the curve length of NCS D. And $L_{straight}$ is length of solid line from 'a' to 'b' in Fig. 2. Fig. 2 shows the NCS D and effective bandwidth.

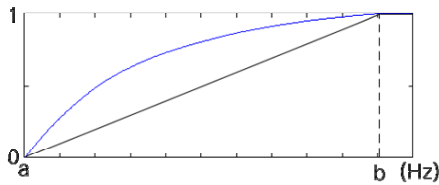


Fig. 2 Spectrum and Normalized Cumulative Spectral Distribution and Effective Bandwidth

Effective bandwidth is decided when $S_{CN}(f)$ is lower than threshold. The threshold is from experimental result. After that NCS D Curvature is extracted on effective bandwidth. k_{CN} is close to 1 when the NCS D curve is similar to solid line. If the NCS D line is curved a lot, k_{CN} will be large. So it can represent the NCS D and relative spectral distribution.

III. MUSIC COMPOSED WITH INSTRUMENTS

To confirm property for music composed by instruments, we classified music into classic by many instruments, classic by one instrument, classic by electronic instrument, new age, ballad, jazz, dance, old pop, pop, rock and techno. Fig. 3 – Fig. 13 show the spectrum and NCS D

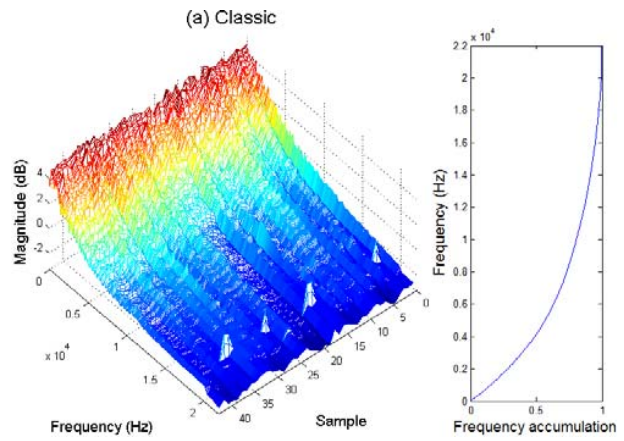


Fig. 3 Spectrum and NCS D (a) Classic by All Instruments

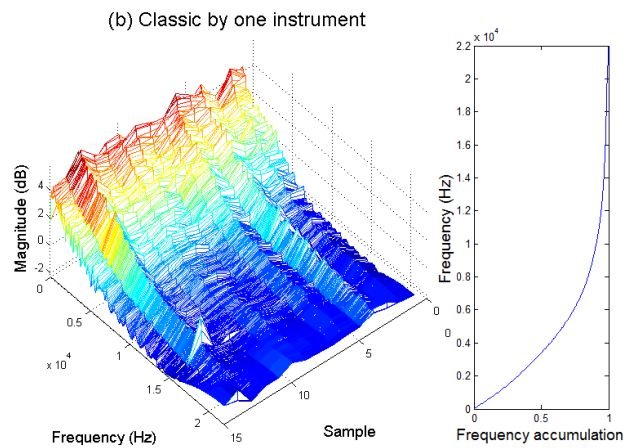


Fig. 4 Spectrum and NCS D (b) Classic by One Instrument

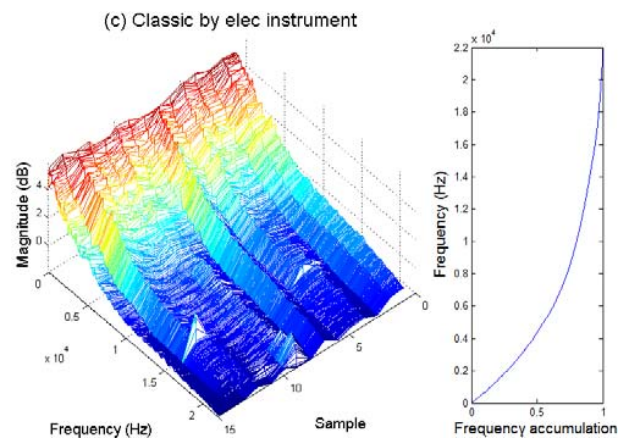


Fig. 5 Spectrum and NCS D (c) Classic by Elec Instruments

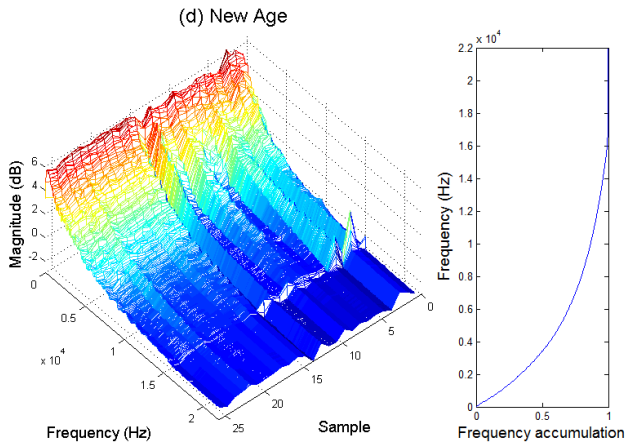


Fig. 6 Spectrum and NCSD (d) New Age

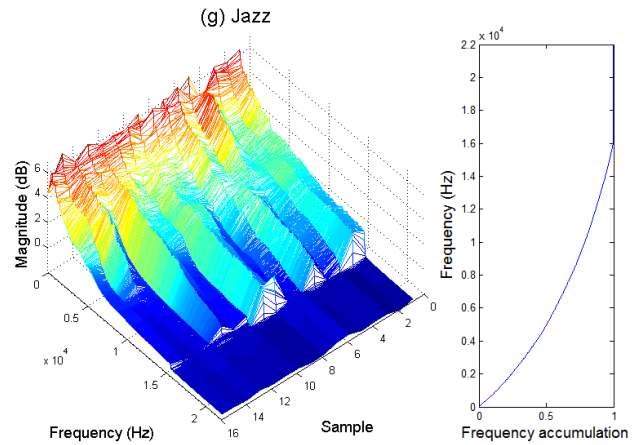


Fig. 9 Spectrum and NCSD (g) Jazz

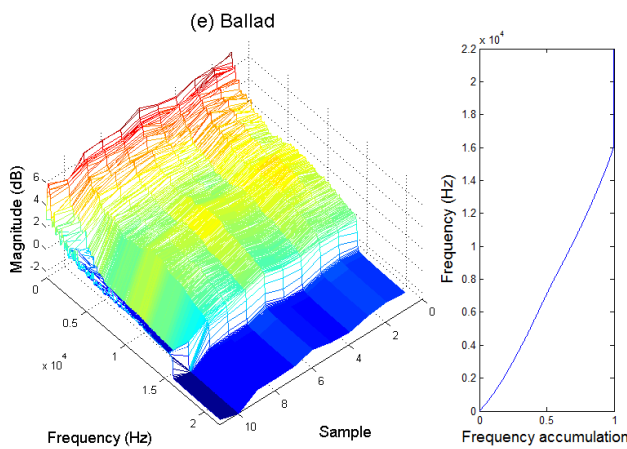


Fig. 7 Spectrum and NCSD (e) Ballad

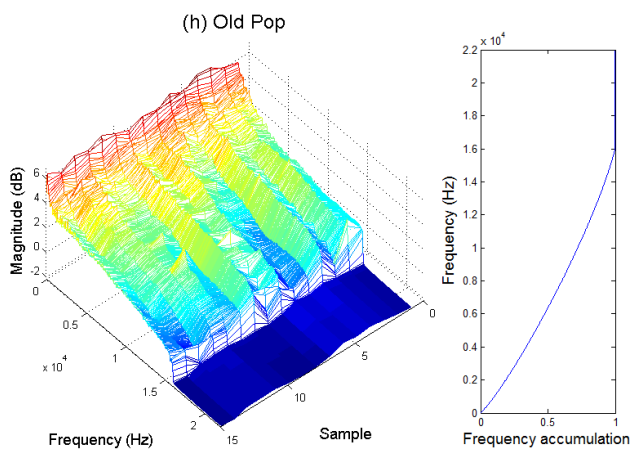


Fig. 10 Spectrum and NCSD (h) Old Pop

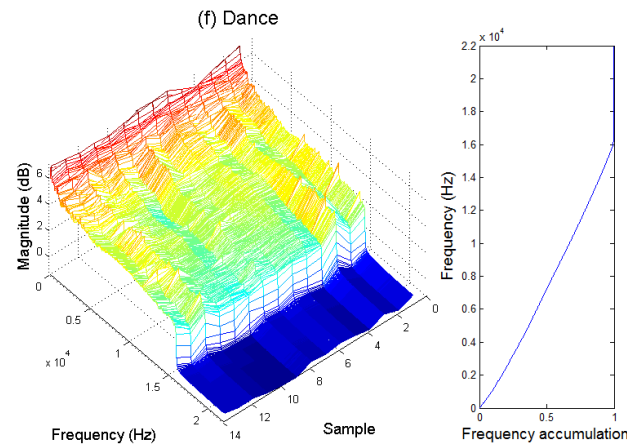


Fig. 8 Spectrum and NCSD (f) Dance Music

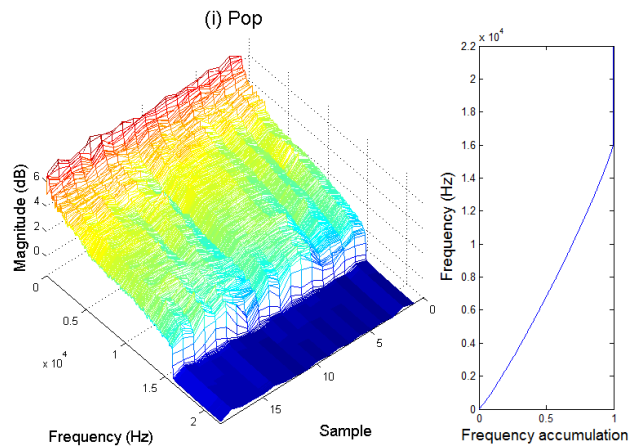


Fig. 11 Spectrum and NCSD (i) Pop

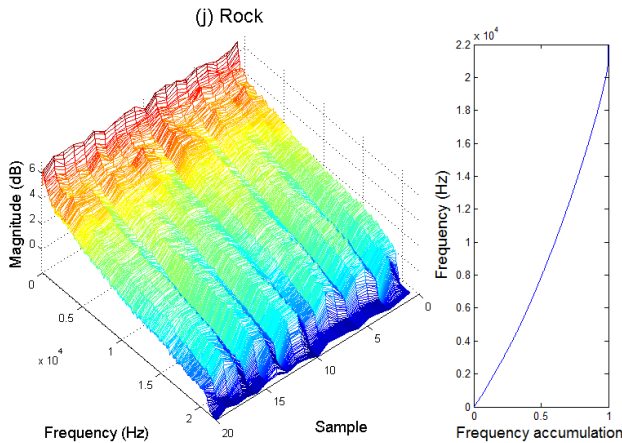


Fig. 12 Spectrum and NCS D (j) Rock

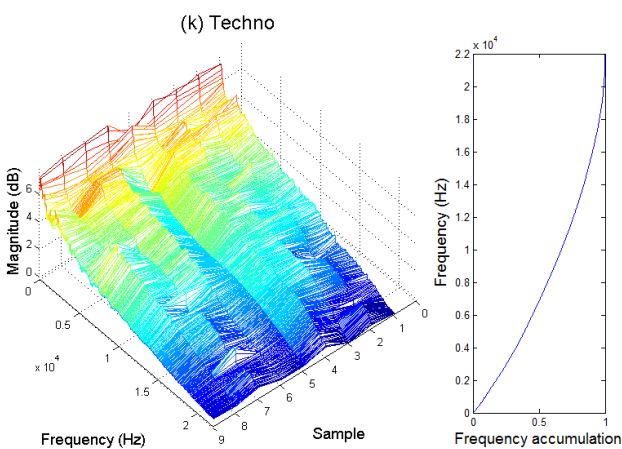


Fig. 13 Spectrum and NCS D (k) Techno

Table I shows k_{CN} for all kinds of music that is classified in this paper. On the table 1, k_{CN} of sacred song that is just made of voice is large. And Classic and New age have large point like sacred song. They did not use voice but they show similar point with sacred song. But ballad, dance, jazz, old pop, pop, rock and techno have small point. Classic and New age give

TABLE I
CURVATURE

Music Genre	k_{CN} (%)
Sacred Song	53.59
Classic by all instruments	71.04
Classic by one instrument	123.87
Classic by one elec instrument	74.40
New Age	73.29
Ballad	7.02
Dance	13.50
Jazz	22.52
Old Pop	8.43
Pop	6.97
Rock	14.37
Techno	31.48

k_{CN} =NCS D Curvature

psychological balance to people. And it is generally qualified to many people for a long time or nowadays.

IV. CONCLUSION

In music perception of cognitive psychological, the curvature of normalized cumulative spectral distribution is high for music which gives psychological balance. That means consonance of sounds in spectrum is well. Commercial music such as pop or techno is composed stimulative and played by lots of musical instruments. In this time, each musical instrument played with each characteristic on the spectrum since each instrument has main part of spectrum different from the others. So it figures like a horizontal line on the spectrum. It makes gradient of normalized cumulative spectral distribution to constant nearly.

We express our emotion through voice. Let assume that voice is one of instrument. In this time, the music which is only composed by voice can effect a change to psychologies. This kind of music is characterized by the output of vocal organ. In this paper we focused about that music which is only composed by voice gives psychological balance. In this kind of music, energy is mainly distributed on the low frequency. Energy is decreased rapidly by the increasing frequency. And the spectrum is distributed for all auditory frequency bandwidth even though it is only composed by voice.

In the music which gives psychological balance, it is established by each sound of instrument for entire harmony, so the energy is mainly distributed on the low frequency. Even though it is composed entirely of a lot of instruments, all instruments played like one instrument without characterizing individually. So magnitude is decreased by increasing frequency. This property is the same as of that music which is only composed of vocal organ of human gives psychological balance. It is shown that the spectral property of music made of voice is revealed at music gives psychological balance regardless of musical composition of instrument.

In this paper, we confirmed the property of music gives psychological balance by the normalized cumulative spectral distribution, and the curvature represented this property on the normalized cumulative spectral distribution.

REFERENCES

- [1] B. Logan and A. Salomon, "A Music Similarity Function based on Signal Analysis," Proceedings of the IEEE International Conference on Multimedia and Expo (ICME2001), pp.745-748, 2001.
- [2] Beth Logan, "Mel Frequency Cepstral Coefficients for music modeling," Proceedings of the First International Symposium on Music Information Retrieval (ISMIR), 2000.
- [3] L. Rabiner and B. H. Juang, Fundamentals of speech recognition, Englewood Cliffs, NJ: Prentice Hall, 1993.
- [4] G. Peeters, "A Large Set of Audio Features for Sound Description (Similarity and Classification) in the Cuidado Project," IRCAM, Tech. Rep., 2004.
- [5] S. Essid, G. Richard, B. David, "Instrument recognition in polyphonic music based on automatic taxonomies," IEEE Trans. on Audio, Speech and Language processing, Vol. 14, No. 1, pp. 68-80, Jan. 2006
- [6] S. Essid, G. Richard, B. David, "musical instrument recognition based on class pairwise feature selection," in Proc. 5th Int. Conf. Music Information Retrieval (ISMIR), Barcelona, Spain, Oct. 2004.

- [7] R. F. Voss and J. Clarke, "1/f noise' in music and speech," Nature, Vol. 258, Issue 5533, pp. 317-318, 1975.
- [8] R. F. Voss and J. Clarke, "'1/f noise" in music: Music from 1/f noise," J. Acoust. Soc. Am., Vol. 63, pp. 258-263, Jan. 1978.



Y.H. Song (M'07) became a Member (M) of **ASK** in 2007. He is from Republic of Korea and was born in 1981. Next, He received the B.S. degree in Electronic Engineering from Soongsil University in 2007. He is currently the under M.S. degree at Soongsil University in Seoul, Korea.

He received the military service from Jun, 2001 to Aug, 2003. He researched "The Identification of Sound Source for Infrasound" with the department of national defense. His research interests include speech signal processing, speech synthesis, speech

recognition, speech coding, and audio coding.

Mr. Song became a Member of ASK in 2007.