

A Comparative Voice Analysis between Original Singer and Mimic Singer in the Speech Signal Processing

Seong-Geon Bae ^{#1}

¹*School of Software Application, Kangnam University, Gyunggido, Korea.*

Sangbum Park²

²*Information and Telecommunication of Department, Soongsil University, Seoul, Korea.
369, Sangdo-ro Dongjak-gu, Seoul*

Myung-Jin Bae ^{*3}

²*Information and Telecommunication of Department, Soongsil University, Seoul, Korea.
369, Sangdo-ro Dongjak-gu, Seoul*

**Correspondence author*

Abstract

Unlike the trotters of the Japanese colonial period, the singer-songwriter used the heavy bass that standard pop singer singers showed. The singing voices of this group led not only to listeners but also to other singers. In this paper, through the analysis of the voice, we analyzed the singing voices of the singers and the singers of the singers' singing triangles "and" foggy jangchungdong park "And MOS test. We studied the singing voice of Baeho by using the frequency, amplitude, and rhythm of sound which are the three elements of sound. In the case of the "returning triangles", the connections are made naturally, with little or no change in tone, as if they were spoken of bass and midrange. In the case of "Foggy Jangchung-dan Park", deep vibrations enter every section of the voice, so when you listen to the song, you can feel the depth. As a result of a comparison study using the MOS TEST, the most similar song to the original sound source is the returning triangle of Na hoon. The foggy Jangchungdong Park, which was sung by Kim Sang-jin, was the second most similarity. Finally, when the original sound source was changed, the sound volume changed, but the rhythm, duration, and amplitude did not change.

Key words: delta, returning delta, vibration, frequency energy distance average, frequency-based energy distance average

INTRODUCTION

The word "trot" in English means "walk fast", "jog at a busy pace", and in 1914, the term "tempo" The fox-trot, which refers to the rhythm of the dance or the style of the dance, is popular since it is popular ^[1]. In Korea, Trout began to develop again in the 1960s, and by the 1970s it was based on Foxtrot's four-fourths of rhythm, but with a distinctive beat-up style, ^{[2] [3] [4]}. At this time, the finished trout is now the trout.

In the first half of the 1960s, the standard pop troupe was revived as a new mainstream.

The earliest works of the early singers, called "Dumé Mountain", are "Trouble" and "Chadie Chicks". The early work was a standard pop-based mix of jazz and Latin music ^[5] ^[6]. His singing is different from trot singers from the Japanese colonial period until the 1950s. He emphasized the profound bass that standard pop singer singers showed with its distinctive vibrations, and used the way that the government gave a loud treble. It also stands at the center of the 1960s-style new male trot system, starting with the singers Gyokgak Taek Taek and the Nam Ilhae, and leading to Nam Jin. Bae Ho is a long-time singer, who has been a popular peak at the age of twenty-nine and has produced fake subordinate recordings that evoke the singer-song singer syndrome for a long time and imitate him for a long time ^[3] ^[4] ^[5] ^[6] ^[9] ^[12] ^[13] ^[14] ^[15] ^[16]. When I listen to singer singing, I feel softness and appeal. The reason is that the song is played across bass and midrange. In addition, the vocal feature allows the listener to feel sympathetic by approaching the song intimately. The singing voices of this group of singers, not only the listeners but also the singers who imitated the singers ^[4] ^[6] ^[7] ^[8] ^[11] ^[12] ^[15]. Therefore, in this paper, through the analysis of the voice, we analyzed the singing voices of the singers of the singers and the other singers who imitated the sound sources of the "triangulation of the returning" and "misty jangchung park" Time, amplitude, etc. The composition of the paper is as follows. In Chapter 2, we study the singing voice characteristics of Baeho using three elements of sound. In Chapter 3, we experiment and result of voice characteristics, and finally in Chapter 4.

SINGING VOICE FEATURES AND ANALYSIS OF BAE HO

Singing voice of Baeho

The characteristics of the voice of Choo called "the returning triangle" studied by using the three elements of sound are singing, and the connection is made naturally as if saying the bass and middle sound with little change of tone. The reason why the voices feel so soft is that they usually sing between 100 and 300 Hz, which is the male voice tone. Park Jin-do, a singer who sings the song of Baeho's song, resembles the song voice, but the difference is the middle-middle dialect.

"Foggy Jangchungdong Park" is one of the popular songs along with "The Returning Triangle" [8][9][10][15][16][17][18]. When you listen to "Foggy Jangchung-dan Park", you can feel deep and listen to the heart because the deep bass of the bass, mid-tune, and treble of the song comes in. Also, the intensity of the song is well processed and sounds natural.

Proposed method

The three elements of the existing sound are frequency, amplitude, and rhythm of the sound, and further classified into the duration of the sound. Frequency is the number of times the radio wave vibrates for one second when moving in space and is determined by the frequency of the sound waves. Amplitude refers to the degree of tremor of vibration or sound, and refers to the amount of energy that travels from the center to the maximum distance or period of time during a period of displacement, when there is periodic vibration [7][11][12][13][12][13].

In this study, we experimented by changing the sound, rhythm, duration, and amplitude using three elements of sound. In addition, the energy distance averaged over the frequency was used to compare the singing voices of the original singers and the singers. In the experiment, each sound source was divided into one each, and the frames were divided every 1 second. In the comparative analysis, the sound sources of the original and the mantissa were compared by two songs, and the measured values were calculated using the energy distance average expression on the frequency [14][15].

$$f_1 = \frac{(e_1^2 + e_2^2 + e_3^2 \dots e_n^2)}{s}$$

$$f_2 = \frac{(e_1^2 + e_2^2 + e_3^2 \dots e_n^2)}{s}$$

$$\vdots$$

$$f_n = \frac{(e_1^2 + e_2^2 + e_3^2 \dots e_n^2)}{s}$$

$$\therefore f_n - f_n = f_{average}$$

F: Frequency average deviation

S: Sampling rate

E: Energy

EXPERIMENTS AND RESULTS

When you listen to the song "Backing Triangle", you will see the voices characteristic of bass softness and appeal. The sound source is evaluated by one of the first sentence of the "Returning Triangle", "One Sighing ~ Lonely ~ Manai ~" and one of the first sentence of "Misty Jangchungdan Park".

We compared five kinds of softness of the bass felt when singing by different mangchang singers that changed the amplitude, soundness, rhythm, and duration of the sound source of "Returning Triangle", "Foggy Jangchungdan Park" TEST. MOS TEST was performed on 30 subjects without distinction between male and female, and the average was shown based on the original sound source (5 points). In addition, the average energy distance of the four frequency sound sources was obtained using the above equation.

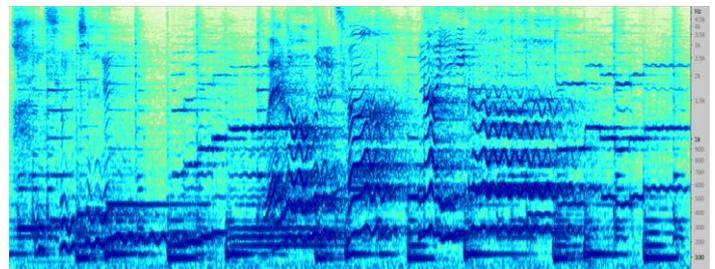


Figure 1: " One paragraph of "Turn Around at Samgakji"
 "Let's breathe ~ Loneliness ~ Man ~"

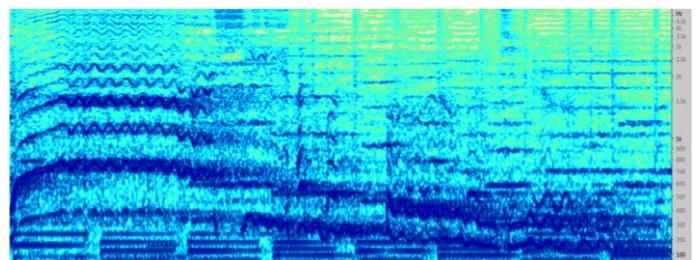


Figure 2: "Foggy Jangchoongdan Stadium" One section
 "Once again~"

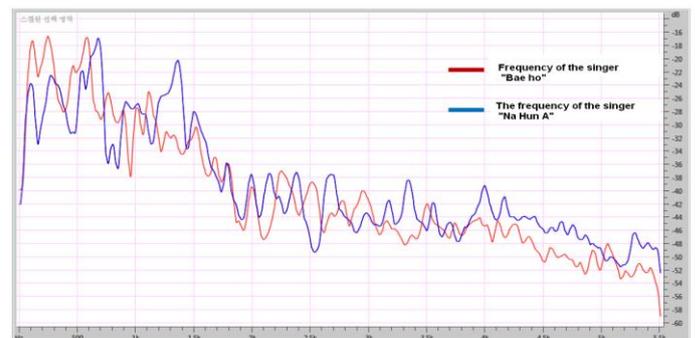


Figure 3: One measure of the "Turn Around at Samgakji",
 "sighs ~ lonely ~ mana" ~ comparison spectrum

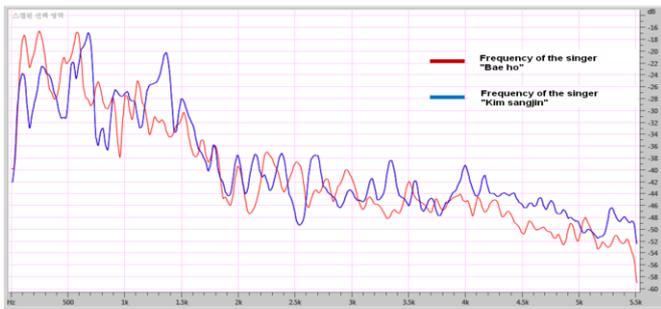


Figure 4: A comparison spectrum of the Foggy Jangchoongdan Stadium One section "Once again ~"



Figure 6: Foggy Jangchoongdan Stadium One section "Once again ~" Original sound source change spectrum

The speech file was sampled at 8000 Hz and 16-bit quantized. Audition CC and Cool Edit Pro 2.1 were used as software. Figure 1 is the original sound source of "a sighing ~ lonely ~ manai ~", a measure of the "returning triangle". Figure 2 is the original sound source of "Once again ~ Touching ~", a measure of "Foggy Jangchung Park". The two original sound sources were compared and analyzed by varying the loudness, rhythm, duration, and amplitude of sound. In addition, it was analyzed and compared with the sound sources which were sung by other singers. As you can see in Figure 3, the phrase "sighs, ~ lonely ~ manaiga ~", which is called by Bae Ho, comes out to be dB lower than the sound that Na na hua calls. Figure 4 also shows that a singing song played by a singer is smoother than a singing by a singer Kim Hyung Joong.

Figure 5 shows the spectrum of the original sound source of "the sighing ~ lonely ~ manaiga ~", which is a measure of the "returning triangle", changing the loudness, rhythm, duration, and amplitude. Figure 6 shows the spectrum of the original sound source of "once again touched ~" which is a measure of "Foggy Jangchung-dan Park". In Figures 5 and 6, the frequency difference is not noticeable in the original sound, but the difference is large when the sound volume is changed.

Table 1 shows the energy distance averages over the frequency of the returning triangles. The energy distance averaged over the frequency range is derived from the above formula. The frequency domain energy was divided into 21 segments by 1 second. Figure 7 shows a graph of energy distance averages over the frequency of the returning triangles. The graph shows the average energy distance of 21 segments. The analysis result shows that the energy range of 0 to 11 seconds is relatively larger than the returning triangulation. When we sang the triangles, we can see that the energy distance average is relatively higher than that of Na Huna in 2 ~ 4 sec interval and 10 ~ 11 sec interval.

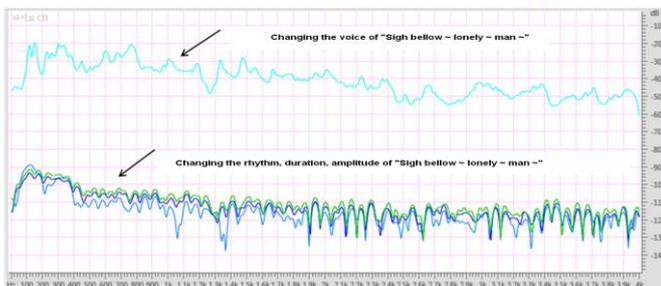


Figure 5: One measure of the "Turn Around at Samgakji", "sighs ~ lonely ~ mana~" Original sound source change spectrum

Table 1: Frequency-based energy distance average of Turn Around at Samgakji

Singer\sec	0~1	0.5~1.5	1~2	1.5~2.5	2~3	2.5~3.5	3~4	3.5~ 4.5	4~5	4.5~5.5	5~6
Bae	1682	1745	1876	2430	3115	3881	2949	2893	2641	2039	1912
Na	1940	1881	2050	2264	2211	2223	2317	2429	1790	1500	1518
Singer\sec	5.5~6.5	6~7	6.5~7.5	7~8	7.5~8.5	8~9	8.5~ 9.5	9~10	9.5~10.5	10~11	
Bae	1743	1629	2044	2275	1849	1700	1863	2111	3158	2124	
Na	1623	1780	1666	1638	1714	1800	1873	1997	1900	1851	

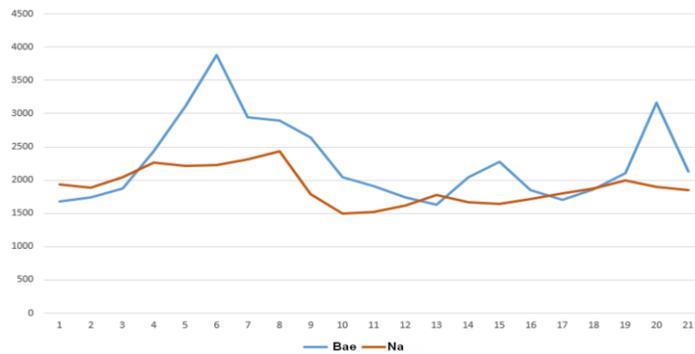


Figure 7: Frequency-based energy distance average graph of Turn Around at Samgakji

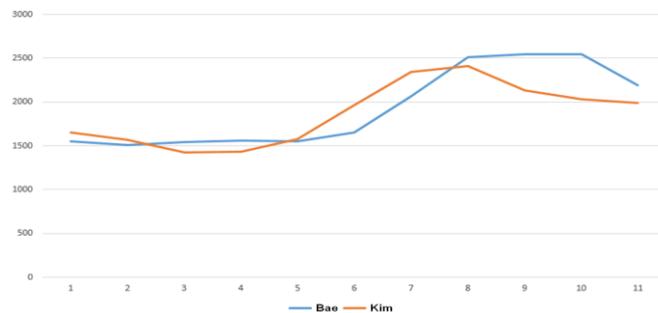


Figure 8: Frequency-based energy distance average graph of Foggy Jangchoongdan Stadium

Table 2: Frequency-based energy distance average of Foggy Jangchoongdan Stadium

Singer\sec	0~1	0.5~1.5	1~2	1.5~2.5	2~3	2.5~3.5	3~4	3.5~4.5	4~5	4.5~5.5	5~6
Bae	1555	1511	1539	1560	1553	1654	2065	2515	2548	2545	2192
Kim	1649	1572	1429	1432	1581	1964	2348	2411	2134	2029	1988

Table 2 shows the frequency distance energy distance of Singer Bae Ho and Kim Sang-jin who fogged Jungchung-dong Park. The average energy distance between Bae Ho and Kim Sang - jin in the foggy Jangchung - dong Park showed no significant difference between 1 ~ 4 seconds and about 400 ~ 5 ~ 6 seconds. Figure 8 shows a graph of energy distance over the frequency of foggy Jangchung-dong park. As shown in the graph, it can be seen that there is little difference in energy distance average over frequency.

Table 1 shows the results of the MOS TEST, which shows the mean values of "returning triangle" and "foggy Jangchung Park" by each sound source. The two original sound sources were classified into five groups according to A (a triangle returning from Namhaea), B (Change of sound source), D (change of rhythm of sound source), E (change of sound source duration), F (change of amplitude of sound source) It is the foggy Jangchunggan park which Kim Sang-jin called. It is the foggy Jangchung park which Kim Sang-jin called the next. When the sound quality of the two sound sources was changed, the average was over (3.5 points). As a result of evaluating the bass softness and appeal when changing the rhythm of two sound sources, it was 1.3. The result of changing the duration of sound (3.3 points) appeared. Finally, the result of changing the amplitude of the two sound sources (3.4 points) came out. As a result of the above MOS TEST, if you change the amplitude of the sound, you can feel the softness and the appeal of the bass more than the higher the sound is, the higher the response is. I got the worst averages. When I changed the rhythm of the sound, I could not feel the softness and the appeal of the bass sounds above average. When the duration of the sound was changed, it was

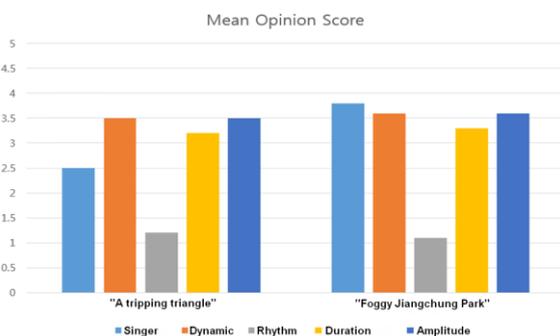


Figure 9: Similar to original sound by each sound source MOS

reproduced 1.5 times slower than the original sound source, and the bass was more emphasized and the highest average was obtained.

CONCLUSION

In general, trot is a word translated from English, which means to walk quickly or to jog with a busy pace. Barrow, a trot singer, produced a profound bass that was shown by standard pop singer singers unlike the trot singers until the 1950s. Baeho emphasized the peculiar vibrations of the singing voice, and used the lyric treble in the temple government. The singing voices of this group led not only to listeners but also to other singers. In this paper, we have obtained the mean energy distance of the other singers who are playing the sound sources of "Returning Triangle" and "Foggy Jangchungdong Park" among the various representative songs of singer 's choir. We compared the song voice with the MOS TEST by changing the loudness, rhythm, duration, and amplitude. As a result, based on the two original sound sources, "the returning triangle" called by Na na Aa appeared similar to the original sound source. Next came Kim Jong - jung 's "Foggy Jangchung - dong Park". When the sound volume is changed, the softness of the bass sound is not more remarkable than that of the original sound source. When I changed the rhythm, the entire song progressed quickly and I could not feel the unique vibrations and softness of the bass. If you change the duration, it progresses slowly and you feel more bass and vibrations so it feels soft and appealing. Finally, when the amplitude was changed, the sound became louder than the original sound source, so it did not sound more appealing than the original sound source.

REFERENCES

- [1] Young-Mee Lee, Trends and Issues in Trot Controversies, The Criticism and Theory Society of Korea. Criticism and Theory. (2008), Vo1.13, No.1, pp.333-68.
- [2] Chan-Joong Jung, Myung-Jin Bae, "A Study on a Vocal Characteristic of Lee Mi-Ja's Song", The Acoustic Society of Korea, Vol.28, No.1, pp.41-42, (2009)
- [3] Jeong-Ho Choi, Chan-Joong Jung, Myung-Jin Bae, "A Study on the Reason Why are Attracted to the Trot", The Acoustic Society of Korea, Vol.29, No.2, pp.414-415, (2010)
- [4] Seong-Geon Bae and Myung-Jin Bae, "Using High Frequency Emphasis as a Personal Speech Bandwidth in Speech Signal," Vol. 17, No. 12/B, pp.6537-6542, 2015.
- [5] Chang-Joong Jung, Seong-Geon Bae, Myung-Sook Kim and Myung-Jin Bae, "Speech Sobriety Test Based on Formant Energy Distribution," IJMUE Vol.8, No.6, pp.209-216, 2013.
- [6] Won-Hee Lee, Seong-Geon Bae and Myung-Jin Bae, "A Study on Improving the Overloaded Speech Waveform to Distinguish Alcohol Intoxication using Spectral Compensation," International Journal of Engineering and Technology, ISSN:0975-4024, Vol.7, No.5, pp.1957-1964, Oct-Nov 2015.
- [7] Seong-Geon Bae, Myung-Sook Kim, and Myung-Jin Bae, "A New Hybrid Non-Uniform Coding with Low Bit Rate for Sound Signal in Near Field Communication," International Conference on ICT Convergence, Proceedings of IEEE ICTC2013, vol., No., pp.1039-1040, Jeju Island Korea, October 14, 2013.
- [8] Jae-Yeon Yoo, "A Study of Acoustic Analyses of Singer's Voice", Graduate School, Taegu University, doctorate thesis, December (2013)
- [9] Woo-Seok Yang "A Study on Historical Transition of Korean Trot Music", Music and ethnicity, Ethnic Music Society, Vol.30, No.0, pp.411-432, (2005)
- [10] Seong-Geon Bae, Myung-Jin Bae "A Study on the effect of heartbeat on the Trot", The Acoustic Society of Korea, pp.162-163, November (2011)
- [11] Hyung-Woo Park, Sang-Hwi Jee, Myung-Jin Bae, "Study on the Confidence-Parameter Estimation through Speech Signal", Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology, Vol.6, No.7, July (2016), pp.101-108
- [12] Seong-Geon Bae, Won-Hee Lee and Myung-Jin Bae, "A Study on Personalized Frequency Bandwidth using Differential Signal in Speech Signal," Proceedings of Symposium of the Korean Institute of communications and Information Sciences, , pp. 868~869, Jan, 2014.
- [13] Seong-geon Bae, KyoungHwa Do and Myung-Jin Bae, "A Study on Personalized Frequency Bandwidth of Speech Signal using Formant to LPC," Proceedings of Symposium of the Korean Institute of communications and Information Sciences, , pp. 669~670, Jan, 2013.
- [14] Chan-Joong Jung, Seong-Young Choi and Myung-Jin Bae, "A Study on Acoustic Analysis of Hook Song," The Journal of The Korean Institute of Communication Sciences, Vol. 35, No. 2, pp. 24~29, Feb, 2010.
- [15] Geum-Ran Baek and Myung-Jin Bae, "A Study on Improving Pitch Search by Varying the number of Subframes for Vocoder," Journal of the Institute of Electronics and Information Engineers, Vol. 49, No. 10, pp. 83~88, Oct, 2012.
- [16] Seong-geon Bae, Myung-Jin Bae and Chang-dong Lee, "A Study on a Drinking Judgement Method of Speech

Signal using the Formant Deviation in the Linear Prediction Coefficient," Proceedings of Symposium of the Korean Institute of communications and Information Sciences, , pp. 667~668, Jan, 2013.

- [17] Wang-Rae Jo, Seong-Young Choi and Myung-Jin Bae, "A Study on the Pitch Search Time Reduction of Cepstrum Method," The Journal of The Korean Institute of Communication Sciences, Vol. 34, No. 10, pp. 310~316, Oct, 2009.
- [18] Wang-Rae Jo, Seong-Young Choi and Myung-Jin Bae, "A Study on the Pitch Search Time Reduction of G.723.1 Vocoder by Improved Hybrid Domain Cross-correlation," The transactions of The Korean Institute of Electrical Engineers, Vol. 59, No. 12, pp. 2324~2328, Dec, 2010.