JESK J Ergon Soc Korea 2019; 38(1): 39-47
http://dx.doi.org/10.5143/JESK.2019.38.1.39
http://jesk.or.kr eISSN:2093-8462

Sound Quality Analysis of Baby Crying Voices

Hyunsu Kim

Dong-eui University, Department of Automotive Engineering, Busan, 47340

아이 울음소리 음질 분석

김 현 수

동의대학교 기계자동차로봇부품공학부

Corresponding Author
Hyunsu Kim
Dong-eui University, Department of Automotive Engineering, Busan, 47340
Mobile : +82-10-3305-9230
Email : hkim7@deu.ac.kr

Received : January 08, 2019
Revised : January 12, 2019
Accepted : January 18, 2019

Copyright@2019 by Ergonomics Society of Korea. All right reserved.

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Objective: This study aims the sound quality analysis of baby crying sound to understand the sound characteristics of baby crying.

Background: Baby crying may be the fundamental communication method between babies and adults. Thus, baby crying can represent various emotion, request, condition, or alarm. Due to those features, a number of studies are reported that medical diagnosis based on baby crying analysis are possible. Recently, unfortunate accident that baby is left in a car leading to sometimes death is often reported, so automatic detection of baby crying is also tried. However, there is few study showing the time-frequency analysis of baby crying which illustrates the intuitive understanding the characteristics of baby crying. Therefore, clear demonstration for the sound analysis for baby crying needs to be investigated.

Method: 10 baby crying sounds are used for time-frequency analysis. For Fast Fourier Transform (FFT), Hann window is utilized, and 50% overlap is applied. Fundamental frequency of baby crying is found, and the magnitudes of harmonics are detected and tabled. Also, different crying patterns are compared with a typical crying sound with the same baby. Lastly, various sound sources, man and woman's voice and an instrument, are investigated and compared.

Results: Baby crying has certain characteristics which show the 1st fundamental component around below 500Hz, and 1~3kHz components tend to be amplified. In addition, different crying patterns show distinguishable sound characteristics.

Conclusion: Baby crying is unique in terms of the sound characteristics which show human ears sensitivity.

Application: The results can assist the study in the field of medical application for diagnosis of baby conditions or baby safety method detecting baby crying sound.

Keywords: Baby crying, Time-frequency analysis, Human ear sensitivity, Harmonics

1. Introduction

어린 아기의 울음은 아기가 가지고 있는 몇 안 되는 의사표현 중 하나이며, 특히 부모의 감성을 자극하는 특수한 소리의 형태를 가지고 있다. 아기의 울음소리가 들릴 때 사람

들은 보통 민감하게 반응을 하며 어떤 경우에는 아이의 울음소리를 매우 거슬리게 느끼기도 한다. 아기의 울음소리만을 듣고 100% 정확한 상태를 말하기 어려운 부분이 있으나 아기의 울음에는 분명 아이의 감정이나 상태에 따라 구분이 가능한 면이 있다. 특히, 아이가 거칠게 울거나 흔히 표현하는 숨이 넘어가는 울음으로 어른을 당황하게 하기도 하고, 마음을 아프게 하거나 슬프게 하는 울음소리의 형태도 존재한다.

아이 울음소리를 주파수 분석을 했을 때 기저 주파수(fundamental frequency)는 약 250~450Hz로 알려져 있으며 아기 울음소리 분석에 대한 기술적 접근은 대략 1960~80년대에 이뤄졌다(LaGasse et al., 2005). 이 당시의 연구는 대부분 아이의 의료적 상태 - 미숙아, 약물중독자 아이 etc. - 와 아이의 울음소리에 대한 상관성을 보기 위한 접근 등의 연구가 이뤄졌고(LaGasse et al., 2005) 이런 접근 방식은 최근까지도 시도되고 있다(Mahmoudian et al., 2018). 2000년대에 신호처리 기법과 인공지능 분석이 발달되면서 아기의 울음소리를 분석하고 인지하는 기술적 접근에 대한 노력이 많아졌다(Saraswathy et al., 2012). Saraswathy et al. (2012)에 의하면 이 당시 논문들은 주로 시간-주파수 영역 분석을 기반으로 인공지능 알고리즘을 활용한 아기 울음소리 감지(detection)에 대한 내용이 많은데, 어떤 분석 기법이 지배적으로 사용되고 있지는 않다.

가장 많이 쓰이는 기법은 음성 인식 기법 중 하나인 MFC (Mel-frequency cepstrum) 기법이다(Muda et al., 2010). MFC 방식은 Cohen and Lavner, 2012; Lavner et al., 2016; Torres et al., 2017에 의해 쓰였고, Cohen and Lavner (2012)은 아이 울음 감지를 위해 인공신경망 알고리즘 중 하나인 k-NN (k-nearest Neural Network)을 이용하였고, Lavner et al. (2016)와 Torres et al. (2017)는 CNN (Convolutional Neural Network) 기법을 활용하였다. 이런 논문들은 시간-주파수 분석에서 아이 울음소리에 대한 직관적이고 명확한 모델링을 만들기보다는 딥러닝을 이용한 감지 그 자체에 더 많은 초점이 맞춰져 있으며 정작 아기 울음소리의 특징을 직관적으로 구분해 내는 분석은 충분히 이뤄지지 않았다.

국내에서는 아이 울음소리 분석에 대한 연구가 매우 부족한 상태인데, 울음소리 주파수 대역폭 분석을 통한 소아호흡기 질환 연구나(Kim and Cho, 2008) 성인과 유아의 음성 인식 비교(Yoo and Lee, 2011) 등의 연구 방향이 대부분이다. 따라서, 크게 아이 울음소리에 대한 연구 방향을 보면 하나는 아이 울음소리를 감지하기 위한 인공지능을 이용한 신호처리이거나, 음성 분석 프로그램을 이용한 질환 진단 같은 의료 분야에서 널리 연구되고 있다. 하지만, 아이 울음소리에 대한 좀 더 직관적인 특징을 알 수 있는 시간-주파수 분석에 대해서는 명확한 특징을 제시하는 연구는 미흡한 상태이다. 따라서, 본 논문에서는 아이 울음소리에 대한 시간-주파수 분석을 통해, 아이 울음소리가 가지는 소리의 특성을 분석하여 아이 울음소리를 이용한 감지나 의료 분야에서 활용하기 용이하도록 하는데 그 목적이 있다.

2. Method

아이 울음소리의 음향학적 특성을 파악하기 위해 FFT (Fast Fourier Transform)을 기반으로 시간-주파수 분석을 진행하였다. 아이 울음의 음원을 직접 측정하기 보다는 향후 빅데이터 분석까지 용이하게 하기 위해 유튜브에서 아이 울음소리의 음원을 가지고 분석하였다. 총 10명의 울음소리를 기본으로 분석하여 분석하였고, 분석 소프트웨어로는 Head Acoustics 사의 ArtemiS Suite을 사용하였다. 음원의 샘플링은 44,100Hz이며, FFT를 할 때 Hann 윈도우를 적용하였고, 시간에 따른 중심 주파수 이동을 위해서는 50%의 오버랩을 사용하였다. 충분한 샘플링(44,100Hz)으로 인해 10,000Hz 이상의 데이터가 분석 가능하지만, 여기서는 주요 성분이 많이 포함되어 있는 5,000Hz 영역까지 분석을 중점적으로 하였다. 아이 울음의 음향 특성을 있는 그대로 분석하기 위해 음향 분석에서 주로 사용되는 A-Weighting(인간의 소리 민감 주파수 가중치)은 적용하지 않았다. ArtemiS 소프트웨어에서 시간-주파수 분해능(resolution)을 결정하기 위한 입력값인 스펙트럼 사이즈는 1024를 사용하였다. 끝으로, 음원은 스테레오이나 좌우 채널의 차이가 거의 없고 좌우 채널의 큰 의미는 없으므로 단일(Mono-Left) 채널만 분석하였다.

3. Results & Discussion

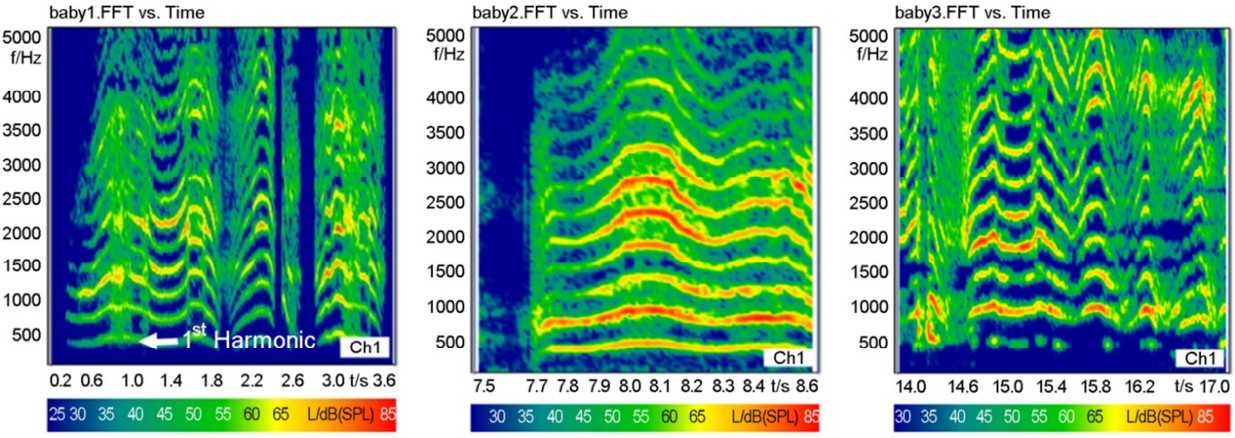
3.1 Time-frequency analysis for 10 babies' crying sound

Figure 1은 3명의 아이(Baby1~3)의 음원을 분석한 시간-주파수 그래프이다. X-축(가로축)은 시간(초)이고 Y-축(세로축)은 주파수(Hz)이다.

그래프의 색깔은 그 시간-주파수에서의 음원의 크기(Magnitude)를 나타낸다. 파란 계열색은 크기가 작고 붉은 계열색은 큰 소리 성분임을 보여준다. 시간 영역을 보면 대략 수 초 간격(2~5초) 정도의 울음소리를 분석한 것을 볼 수 있다.

울음소리를 시간-주파수 영역에서 분석한 그래프를 보면 우선 보이는 특징은 기존의 연구에서 알려진 바와 같이 기본 주파수가 약 300~500Hz 대역에서 보여진다. Figure 1에서 보면 가장 낮은 주파수 영역에서 하나의 음원 주파수가 약 500Hz 미만에서 하나의 선을 나타내고 있다. 이것은 아이의 울음소리의 음원(성대에서 발생된 소리) 성분으로 1차 하모닉 성분이다. 이것이 1차 하모닉 성분임을 보여주는 것은 그 이유에 하모닉 성분의 2배수, 3배수 등으로 이어지는 n차 하모닉 성분들이 거의 정확하게 1차 하모닉 성분의 n 배수로 나타나기 때문이다. 이렇게 n차 하모닉 성분으로 나타나는 것은 보통의 음성 신호를 분석할 때 나타나는 것과 비슷하다고 할 수 있다.

아이 울음소리의 큰 특징 중 또 다른 하나는 아이의 울음소리가 약 1kHz 대역에서 약 3kHz 대역 미만 사이에서 큰 특징을 갖는다는 것이다. 이는 대략적으로 아이 울음소리의 기저 주파수의 약 2~3배에서 시작되는데, 특이한 것은 몇 차 하모닉이냐가 아니라 거의 정확하게 1kHz 대역에서 첫 번째 공진(resonance)를 가진 다는 것이다. 예를 들면 기저 주파수가 약 450Hz 대역이면 2차 하모닉인 약 900Hz 대역에서 공진을 가지고 만약 기저 주파수가 300Hz 대역이면 3차 하모닉인 900Hz 대역에서 공진을 갖는다는 것이다. 이것은 아이의 울음소리는 내는 인간의 신체 구조상에 약 1kHz 대역의 소리를 증폭하여 내는 구조를 가지고 있다는 것으로 볼 수 있다. 예를 들면, 길이 L (m)의 파이프에 성대라고 모델링 될 수 있는 스피커가 있다고 하자. 이 스피커에서는 약 300~500Hz의 기본 주파수 성분과 그 하모닉 성분을 내고 있다고 한다면 약 1kHz 대역에 공진 주파수를 갖기 위해서는 f = 음속/(2L)의 관계식에 의해 약 17cm의 파이프 길이가 필요하다. 아이의 신체 사이즈를 고려할 때, 성대(vocal cords)로부터 입 위치까지의 길이는 17cm 보다 작을 것으로 보인다. 따라서, 단순 파이프 모델보다는 구강의 공간과 기도의 길이를 다 고려한 헬름홀쯔(Helmholtz) 모델이 더 적합해 보이며(Kinsler et al., 1999) 정확한 아이의 인체 구조와 소리 증폭과의 상관성에 대한 연구는 추가적으로 더 필요할 것이다.



(a) (b) (c)

Figure 1. Time-Frequency analysis of baby crying sound: (a) baby1, (b) baby2, (c) baby3

Figure 1에서 관찰되는 특징 중 흥미로운 것은 1~3kHz 대역에서 음압레벨이 커지게 되는 영역이 거의 반드시 나타난다는 점이다. 이 대역은 일반 성인 기준으로 가장 귀가 민감하게 반응하는 영역으로 인간이 인지하는 음향을 분석할 때 이 주파수 영역에 더 많은 음압레벨(sound level)을 보정하는 영역이며 이런 가중치를 Table 1에 나타난 A-weighting이라고 한다. 이 표에서 보면 1~4kHz 대역이 가장 사람이 인지하는데 있어 민감하게 반응하는 주파수 대역임을 알 수 있다. 예를 들면, 500Hz 대역보다 1,000Hz 대역을 같은 음압레벨이라면 약 2배 더 잘 들리는 것(약 3.2dB 가중치 차이)이라고 할 수 있는 것이다. 이는 인간의 귀가 아이의 울음소리에 가장 민감하게 반응하도록 발달되어 있는 것이라고 볼 수 있다. 사람의 귀가 1~4kHz 대역의 소리에 민감하다는 것은 잘 알려져 있지만, 이것이 아이의 울음소리가 증폭되는 주파수의 영역과 겹친다는 매우 흥미로운 결과를 보여주는 것이다.

Table 1. Corrections to be added to octave-band levels to convert to A-weighted band levels (Kinsler et al., 1999)

Center frequency (Hz)	Correction (dB)
31.5	-39.4
63	-26.2
125	-16.1
250	-8.6
500	-3.2
1,000	0
2,000	1.2
4,000	1.0
8,000	-1.1

Table 2는 10명의 아이의 울음소리를 Figure 1과 같은 형식으로 모두 분석한 후 각 기저 주파수와 그 주파수의 1차부터 7차 하모닉 성분까지의 크기를 하나의 표로 정리한 것이다. 우선 주목할 것은 울음소리의 대부분은 기저 주파수가 500Hz 미만이지만 가끔은 500Hz 이상의 기저 주파수 성분을 갖는 울음소리도 존재한다는 것이다(Baby6와 9 참조). 이 표에서는 아이 울음소리를 일정한 테스트 조건에서 측정한 것이 아니기 때문에 기저 주파수 성분의 크기를 보면 30~70dB 대에 이르기까지 매우 넓게 분포되어 있는데 이 연구에서는 측정방법에 상관 없이 아이의 울음소리 특성에 관심이 있기 때문에 이 부분은 무시하도록 한다. 또한, Table 2에 분석되어 있는 10명의 아이의 울음소리의 패턴은 악쓰는 울음이나 흐느끼는 울음이 아닌 일반적 울음의 형태이다.

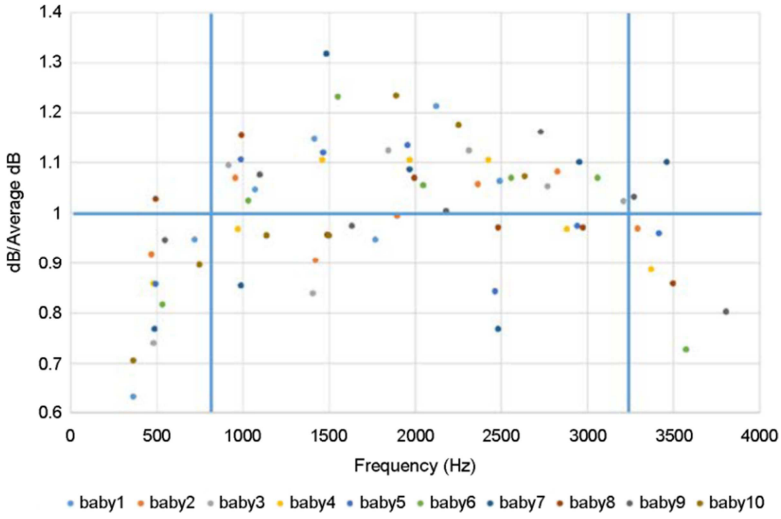
Table 2. Magnitude of order components in dB for 10 babies' crying

	1st order frequency (Hz)	Magnitude of order components (dB)						
		f1	f2	f3	f4	f5	f6	f7
Baby1	363	38	57	63	69	57	73	64
Baby2	469	73	85	72	79	84	86	77
Baby3	480	52	77	59	79	79	74	72
Baby4	480	63	71	81	81	81	71	65
Baby5	491	59	76	77	78	58	67	66
Baby6	530	55	69	83	71	72	72	49
Baby7	486	53	59	91	75	53	76	76
Baby8	491	73	82	68	76	69	69	61
Baby9	547	66	75	68	70	81	72	56
Baby10	363	48	61	65	65	84	80	73

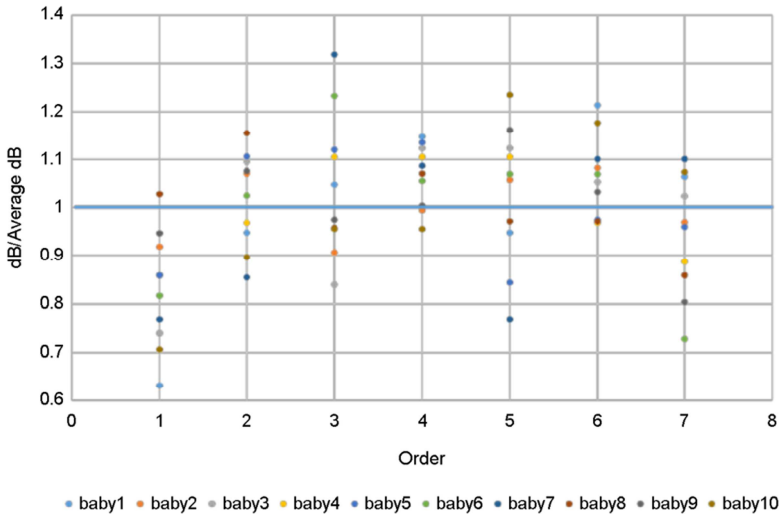
Figure 2(a)는 Table 2에 나와 있는 각 아이 울음소리의 음압레벨 크기 성분을 그 아이 울음소리의 음압레벨 평균값으로 나누어 정량화(normalize)한 값을 주파수의 함수로 표현한 것이다. 주로 1차 하모닉 성분이 있는 500Hz 대역의 성분들은 그 아이의 울음소리의 평균값으로 나눴을 때, 대부분 1보다 작은 값을 갖는 것을 볼 수 있다. 또한 약 3,300Hz가 넘어가면서 아이의 울음 성분은 평균값보다 작은

값들로 구성되고 있음을 알 수 있다. 앞서 묘사한 바와 같이, 1~3kHz 대역은 울음소리의 평균 음압보다 큰, 즉 1보다 값들이 더 많이 분포되어 있는 것을 볼 수 있다. 한 가지 복잡한 것은 단순하게 파이프 모델이나 헬름홀쯔 모델로 표현되기에는 어떤 경우에는 1~3kHz 대역이라도 낮은 음압을 보여주는 영역이 분명 존재한다는 것이다. 이것은 복잡하고 유기적으로 움직이는 인체의 발성 구조상 다양한 소리의 형태가 나타는 것으로 보인다. 예를 들면 성대와 기도를 통해서 입 쪽으로 나오는 소리의 구조와 비강(코) 쪽으로 소리가 전달되고, 혹은 이의 복합적인 형태로 소리가 나가게 될 수 있는 복잡한 구조를 가지고 있는 형태로 인한 것이라고 볼 수 있다는 것이다.

Figure 2(b)는 같은 데이터를 가지고 정량화된 음압을 하모닉 성분의 함수로 다시 편성한 그래프이다. 앞에서 설명된 것처럼, 1차 하모닉 성분은 대부분 작은 음압(1보다 작은 정량화 값)을 가진다. 2차 하모닉은 약 절반정도가 평균값보다 큰 결과를 보여주는데, 이는 1차 하모닉의 기저 주파수가 너무 낮으면 2차 하모닉이 1kHz 보다 너무 낮게 되고 이런 아이 울음소리들이 낮은 2차 하모닉을 갖게 하는



(a)



(b)

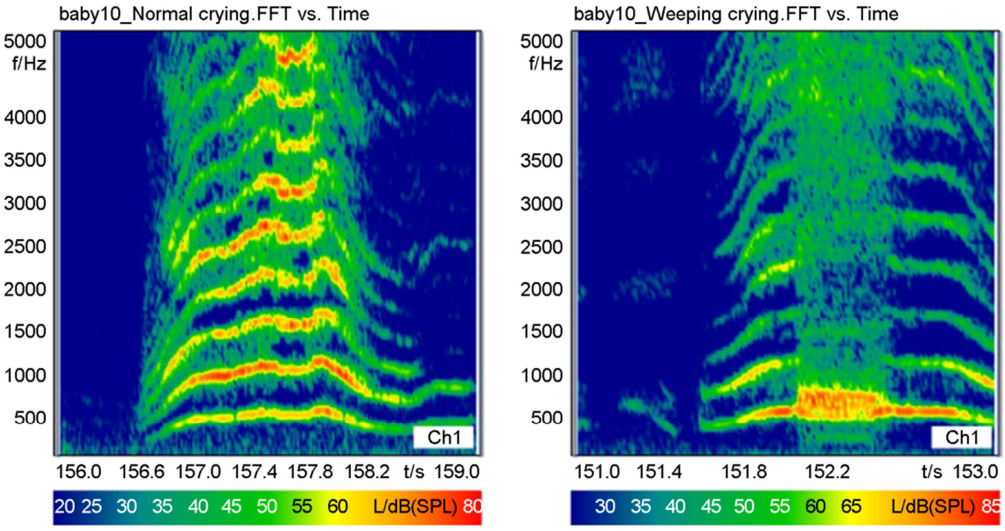
Figure 2. Comparison of normalized magnitudes of 10 babies' crying sound: (a) as a function of frequency, (b) as a function of order

요인이 된다.

이 표에서는 몇 차 하모닉이 아이 울음소리의 지배적이라는 결론을 내릴 수 없다는 것을 보여준다. 다만 아이 울음소리의 기저 주파수가 약 500Hz 미만이며 1~3kHz 대역에서 공진이 나타날 수 있다는 특성을 고려할 때 약 2~6차 하모닉들이 대체적으로 큰 값을 갖는다는 경향만을 볼 수 있다.

3.2 Time-frequency analysis for different crying patterns

Figure 3는 아이 울음소리의 특징이 같은 아이라도 울음의 형태가 다를 때 구분될 수 있다는 것을 보여준다. Figure 3(a)는 한 아이의 보통 울음소리이다. 다른 아이의 보통 울음소리와 마찬가지로 약 500Hz 미만에서 기저 주파수를 가지고 약 1kHz 대역에서 큰 주파수 성분을 가지고 약 1~3kHz 대역에 많은 울음 성분들이 분포되어 있는 것을 볼 수 있다. 하지만, "흐응"하며 흐느끼는 울음소리는 Figure 3(b)에서 보이는 바와 같이 약 500~800Hz 대역의 낮은 주파수 대역에 소리 성분이 집중되어 있는 것을 볼 수 있다. 이는 앞의 보통의 울음소리, 즉 자신의 어떠한 의사표현을 하거나 위험을 알리거나 자신을 일반적으로 알리고자 하는 아이의 울음소리가 성인 귀에 민감하게 잘 들리도록 그 음향 성분이 분포되어 있는 것과 달리, 소위 흐느끼는 울음은 사람 귀에 민감하기 보다는 귀가 민감한 영역보다 낮은 주파수에 분포되어 일반적으로 저주파 일수록 귀보다는 가슴으로 느끼게 하는 사람의 인지 능력에 호소하는 것으로 보이며 매우 흥미로운 음향 분석이라고 할 수 있겠다.

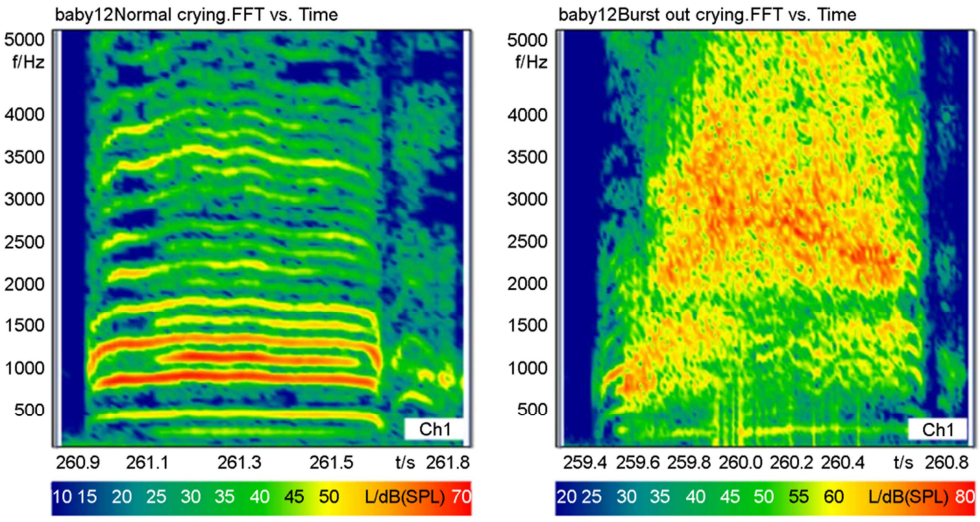


(a) (b)

Figure 3. Comparison of crying sound noise with (a) normal crying and (b) sobbing

Figure 4는 한 아이가 보통 울음소리와 흔히 말하는 악을 쓰며 울 때의 소음의 특성을 비교하였다. Figure 4(a)는 그냥 울 때의 음향 특성으로 앞에서 보여주는 다른 울음소리와 그 특성이 크게 다르지 않다. 하지만, Figure 4(b)에는 같은 아이가 악을 쓰며 울 때 나타나는 음향 특성인데, 우선 기저 주파수 성분이 없어지고, 1~2kHz 때 소음의 성분도 다소 생기지만 2~4kHz 대역의 다소 고주파 소리가 큰 성분을 차지하게 된다. 이는 신기하게도 Table 1에서 민감도 가중치의 2~4kHz 대역에서 다른 어떤 주파수 영역보다는 그 민감도가 가중되는 것을 볼 수 있는데 이 주파수 영역과 일치하는 것이다. 아이가 악을 쓸 때는 위급성을 나타내고자 하는 것일 수 있고, 혹은 본인의 의사를 매우 강력하게 표현하는 것으로 이해할 수 있겠는데, 이러한 표현을 울음으로 나타날 때 가장 사람에게 거슬릴 수 있는 소리의 성분으로 표현을 한다는 것이다. 또한 악을 쓰는 울음 형태는 음향 성분들이 일정한 sine 주파수들의 조합인 하모닉 성분

들로 이루어진 일반적인 울음 패턴과 달리, Figure 4(b)에서 보는 바와 같이 2~4kHz 대역을 모두 채우는 이산(random) 주파수의 성분을 가진다. 보통 자동차에서 나는 엔진음들이 대부분 엔진의 회전수와 연관된 하모닉 성분들이지만 매우 거친 소음이나 기류음 같이 거슬리는 소음들이 이산 주파수 성분들로 표현된다는 것을 상기해 볼 때 매우 흥미로운 분석 결과라고 하겠다.

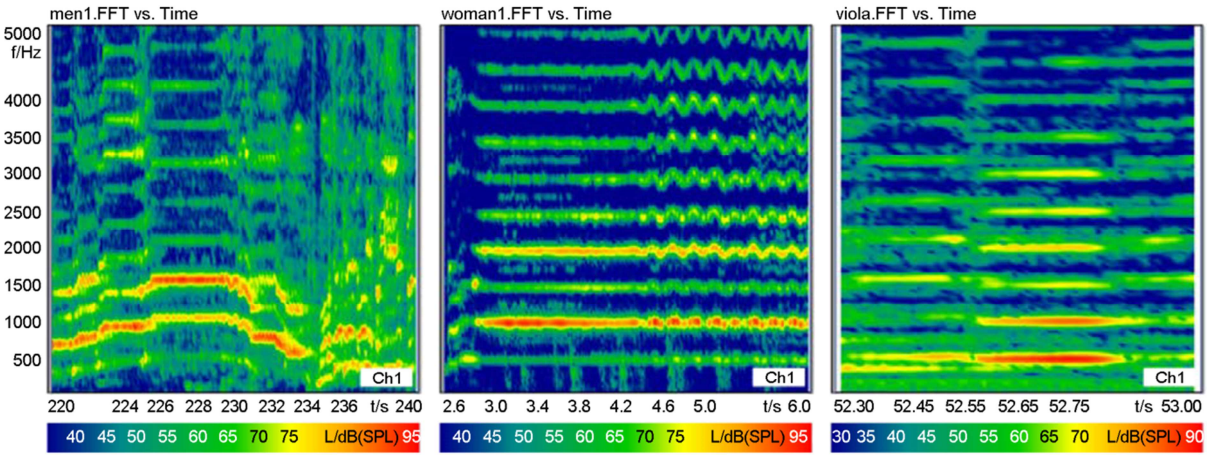


(a) (b)

Figure 4. Comparison of crying sound noise with (a) normal crying and (b) crying with a groan

3.3 Time-frequency analysis for different sound sources

Figure 5는 비슷한 주파수 대역을 가진 다른 소리, 즉 성인 남녀의 목소리와 악기 소리의 특성을 비교하였다. 악기는 약 500Hz 대역 소



(a) (b) (c)

Figure 5. Comparison of other sound with around 500~1,000Hz fundamental frequency with (a) man voice, (b) woman voice, (c) viola instrument

리를 낼 수 있는 여러 악기 중 비올라를 선택하였다. 다양한 악기를 비교 분석하는 것은 향후 추가적인 연구가 필요할 것이다. Figure 5(a)와 (b)는 성인 남녀 목소리를 각각 비교한 것인데, 아이 울음소리와 달리 기본 주파수 성분에서 2차~4차 성분만이 약 2kHz 대역 미만에서 보이고 그 이상의 주파수 성분은 적게 보인다. 악기 소리의 분석인 Figure 5(c)에서는 1차의 기저 주파수가 가장 크고 2차 하모닉 성분이 그 다음으로 커 보인다. 그 이상의 성분은 미미하다. 다만, 이러한 다른 소리에 대한 샘플은 아이 울음소리와 같이 많은 수를 분석한 것은 아니다. 하지만, 이 연구에서는 아이의 울음소리의 특성을 분석하기 위한 목적을 위해 다른 음향 특성을 참고하기 위함으로 추가적인 음향 분석이 필요하겠다.

4. Conclusion

아이의 울음소리는 인간으로서 한 아이가 자신의 의사표현을 하는 가장 원초적이고 부모와 같은 보호자와 소통하는 매우 유용한 수단이다. 아이의 울음소리를 분석하는 것은 의료, 안전 분야에서 널리 사용될 수 있다. 예를 들어 아이의 울음소리의 음향 특성과 울음 패턴을 비교 분석함으로 아이의 상태, 즉 흥분이나 슬픔, 위험 등의 상황을 일차적으로 판단할 수도 있다. 또한, 예를 들어, 기관지 부분에 질병이 있는 경우 아이의 울음 주파수 변경 등을 추적하면서 호전 상태를 파악하는 등 의료 분야에서 아이의 울음소리를 통한 연구를 진행할 수 있겠다. 본 연구에서는 아이의 울음소리에 대한 시간-주파수 분석을 통해 아이의 울음소리가 가지는 기초적이고 직관적인 음향 분석을 하였고, 아이의 울음소리의 특성을 분류하였다. 이에 제시하는 주요한 결론은 다음과 같다.

- 아이의 울음소리는 약 500Hz 미만에서 기저 주파수를 가지며 1~3kHz 대역에서 가장 큰 증폭을 가지는 경향이 있다.
- 아이의 울음소리는 성대에서 울리는 기저 주파수 성분보다는 성대, 기도, 구강, 비강 구조에 따라 소리를 증폭시키는 것으로 보인다.
- 아이의 울음소리는 성인의 귀가 가장 민감하게 받아들이는 A-weighting의 민감한 주파수 영역에 많이 겹치며 이는 부모를 비롯한 보호자인 성인에게 가장 효과적으로 전달되는 음향 특성이다.
- 아이가 흐느낄 때는 귀에 민감하고 거슬리는 1~3kHz 대역보다 1kHz 보다 낮은 영역에서 소리가 형성되어 소위 연민을 느끼게 하는 소리의 성분으로 구성되어 있다.
- 아이가 악을 쓸 때는 사람 귀에 가장 예민한 2~4 Hz 대역에 이산 주파수 성분을 가지며 이는 사람에게 가장 거슬리며 강한 의사표현임을 느끼게 하는 성분들로 구성되어 있다.
- 성인 남녀나 악기의 소리의 특징은 아이 울음소리의 특성과 달리 기저 주파수 성분과 그의 4차 미만의 성분들이 2kHz 미만에 구성되며 그 이상의 소음은 매우 적어지는 특징을 가진다.

본 연구의 음향 분석은 기존의 연구가 인공지능이나 음성 인식 알고리즘을 이용하여 아이의 울음을 구분하는 흡사 블랙박스 같은 분석보다 명확하게 아이의 울음소리의 특성을 보여준다. 이러한 음향 분석이 기초적으로 선행되어야 향후 아이 울음소리를 이용한 감지나 의료 분야 사용에 크게 도움이 될 것이라 본다. 그럼에도 불구하고 향후에 추가적으로 진행되어야 할 연구는 더 많은 아이의 울음소리에 대한 분석과 더 다양한 울음소리의 형태에 따른 음향 분석이 필요하겠다. 또한 성인 소리나 악기, 동물 등과 같은 소리와 구분하기 위한 더 많은 연구가 추가적으로 진행되어야 할 것이다

Acknowledgement

이 논문은 2018학년도 동의대학교 교내연구비에 의해 연구되었음(과제번호 201802840001).

References

Cohen, R. and Yizhar, L.., "Infant Cry Analysis and Detection". In *2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel*, 1-5, 2012. IEEE. doi:10.1109/EEEI.2012.6376996.

Kim, B.H. and Cho, D., "A Study on Infant Respiratory Diseases Diagnosis Using Frequency Bandwidth Analysis of Crying Waveform". *The Journal of Korean Institute of Communications and Information Sciences*, 33(12), 1123-1130, 2008.

Kinsler, L.E., Frey, A.R., Coppens, A.B. and Sanders, J.V., *Fundamentals of Acoustics*. 4th ed. New York: John Wiley and Sons, Inc, 1999.

LaGasse, L.L., A. Rebecca, N. and Lester, M.B., "Assessment of Infant Cry: Acoustic Cry Analysis and Parental Perception". *Mental Retardation and Developmental Disabilities Research Reviews*, 11(1), 83-93, 2005. doi:10.1002/mrdd.20050.

Lavner, Y., Cohen, R., Ruinskiy, D. and Ijzerman, H., "Baby Cry Detection in Domestic Environment Using Deep Learning". In *2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE)*, 1-5, 2016. IEEE. doi:10.1109/ ICSEE.2016.7806117.

Mahmoudian, S., Aminrasouli, N., Ahmadi, Z.Z., Lenarz, T. and Farhadi, M., "Acoustic Analysis of Crying Signal in Infants with Disabling Hearing Impairment". *Journal of Voice*, 2018. doi:10.1016/j.jvoice.2018.05.016.

Muda, L., Begam, M. and Elamvazuthi, I., "Voice Recognition Algorithms Using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques", 2010. http://arxiv.org/abs/1003.4083.

Saraswathy, J., Hariharan, M., Yaacob, S. and Khairunizam, W., "Automatic Classification of Infant Cry: A Review". In *2012 International Conference on Biomedical Engineering (ICoBE)*, 543-548, 2012. IEEE. doi:10.1109/ICoBE.2012.6179077.

Torres, R., Battaglino, D. and Lepauloux, L., "Baby Cry Sound Detection: A Comparison of Hand Crafted Features and Deep Learning Approach". In, 168-179, 2017. doi:10.1007/978-3-319-65172-9_15.

Yoo, J.K. and Lee, K.M., "Comparison of Adult and Child's Speech Recognition of Korean". *Journal of the Korea Contents Association*, 11(5), 138-147, 2011.

Author listings

Hyunsu Kim: hkim7@deu.ac.kr
Highest degree: PhD, Department of Mechanical Engineering, The Ohio State University
Position title: Assistant Professor, Department of Automotive Engineering, Dong-eui University
Areas of interest: Human-perception, Sound and vibration, Acoustics, Human Voices